PCT WORLD INTELLECTUAL PROPERTY ORGANIZATION
International Bureau

INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

(51) International Patent Classification 6 :
C12N 15/12, C07K 14/705, 16/28, G01N 33/50, A61K 38/17, C12N 1/21

A1

(11) International Publication Number: **WO 98/45438**

(43) International Publication Date: 15 October 1998 (15.10.98)

(21) International Application Number: PCT/US98/06959

(22) International Filing Date: 8 April 1998 (08.04.98)

(30) Priority Data:

60/043,447	9 April 1997 (09.04.97)	US
60/046,254	12 May 1997 (12.05.97)	US
08/892,745	15 July 1997 (15.07.97)	US

(63) Related by Continuation (CON) or Continuation-in-Part (CIP) to Earlier Applications

US	08/892,745 (CIP)
Filed on	15 July 1997 (15.07.97)
US	60/046,254 (CIP)
Filed on	12 May 1997 (12.05.97)
US	60/043,447 (CIP)
Filed on	9 April 1997 (09.04.97)

(71) Applicant *(for all designated States except US)*: BETH ISRAEL DEACONESS MEDICAL CENTER [US/US]; 330 Brookline Avenue, Boston, MA 02215 (US).

(72) Inventors; and

(75) Inventors/Applicants *(for US only)*: LOWELL, Bradford, B. [US/US]; 6 Tara Road, Southborough, MA 01772 (US). FLIER, Jeffrey, S. [US/US]; 14 Sylvan Avenue, West Newton, MA 02165 (US).

(74) Agents: GRANAHAN, Patricia et al.; Hamilton, Brook, Smith & Reynolds, P.C., Two Militia Drive, Lexington, MA 02173 (US).

(81) Designated States: AU, CA, JP, US, European patent (AT, BE, CH, CY, DE, DK, ES, FI, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE).

Published
With international search report.
Before the expiration of the time limit for amending the claims and to be republished in the event of the receipt of amendments.

(54) Title: UCP3: AN UNCOUPLING PROTEIN HOMOLOGUE

(57) Abstract

The present invention relates to isolated and/or recombinant nucleic acids which encode a mammalian (e.g., human, mouse) uncoupling protein 3 (UCP3) and an alternative form of UCP3 designated UCP3-short form (UCP3sh). In addition, the present invention relates to nucleic acids which hybridize with the UCP3 nucleic acids described herein and functional portions thereof. Also encompassed by the invention are a nucleic acid construct comprising a nucleic acid which encodes a UCP3 protein and a host cell; a host cell comprising the nucleic acid construct which encodes UCP3; and a method for producing mammalian UCP3 comprising introducing into a host cell the nucleic acid construct which encodes UCP3 whereby the nucleic acid is expressed. The present invention also relates to isolated or recombinantly produced UCP3 protein and functional portions thereof. Also encompassed by the invention are a method of identifying an inhibitor (e.g., antibody) or enhancer of UCP3 expression and/or function, and the use of UCP3 inhibitors and enhancers. The present invention also relates to a method of detecting UCP3 in a sample obtained from a individual.

FOR THE PURPOSES OF INFORMATION ONLY

Codes used to identify States party to the PCT on the front pages of pamphlets publishing international applications under the PCT.

AL	Albania
AM	Armenia
AT	Austria
AU	Australia
AZ	Azerbaijan
BA	Bosnia and Herzegovina
BB	Barbados
BE	Belgium
BF	Burkina Faso
BG	Bulgaria
BJ	Benin
BR	Brazil
BY	Belarus
CA	Canada
CF	Central African Republic
CG	Congo
CH	Switzerland
CI	Côte d'Ivoire
CM	Cameroon
CN	China
CU	Cuba
CZ	Czech Republic
DE	Germany
DK	Denmark
EE	Estonia
ES	Spain
FI	Finland
FR	France
GA	Gabon
GB	United Kingdom
GE	Georgia
GH	Ghana
GN	Guinea
GR	Greece
HU	Hungary
IE	Ireland
IL	Israel
IS	Iceland
IT	Italy
JP	Japan
KE	Kenya
KG	Kyrgyzstan
KP	Democratic People's Republic of Korea
KR	Republic of Korea
KZ	Kazakstan
LC	Saint Lucia
LI	Liechtenstein
LK	Sri Lanka
LR	Liberia
LS	Lesotho
LT	Lithuania
LU	Luxembourg
LV	Latvia
MC	Monaco
MD	Republic of Moldova
MG	Madagascar
MK	The former Yugoslav Republic of Macedonia
ML	Mali
MN	Mongolia
MR	Mauritania
MW	Malawi
MX	Mexico
NE	Niger
NL	Netherlands
NO	Norway
NZ	New Zealand
PL	Poland
PT	Portugal
RO	Romania
RU	Russian Federation
SD	Sudan
SE	Sweden
SG	Singapore
SI	Slovenia
SK	Slovakia
SN	Senegal
SZ	Swaziland
TD	Chad
TG	Togo
TJ	Tajikistan
TM	Turkmenistan
TR	Turkey
TT	Trinidad and Tobago
UA	Ukraine
UG	Uganda
US	United States of America
UZ	Uzbekistan
VN	Viet Nam
YU	Yugoslavia
ZW	Zimbabwe

UCP3: AN UNCOUPLING PROTEIN HOMOLOGUE

RELATED APPLICATIONS

This application is a Continuation-in-Part of Application No. 08/892,745 entitled "UCP3: An Uncoupling Protein Homologue Expressed Selectively and Abundantly in Skeletal Muscle and Brown Adipose Tissue" filed July 15, 1997 and claims benefit of U.S. provisional application number 60/046,254 entitled "Discovery of an Alternative Form of UCP3, Designated UCP3-Short Form (UCPsh)" filed May 12, 1997 and U.S. provisional application number 60/043,447, entitled "An Uncoupling Protein Homologue Expressed Selectively and Abundantly in Skeletal Muscle and Brown Adipose Tissue", filed April 9, 1997. The teachings of Application No. 08/892,745, U.S. provisional application number 60/043,447 and U.S. provisional application number 60/046,254 are incorporated herein by reference in their entirety.

GOVERNMENT FUNDING

This work was supported in part by the National Institutes of Health Grants DK02119 and DK49569. Therefore, the U.S. Government has certain rights in the invention.

BACKGROUND

Calories are expended by mitochondria in a highly regulated fashion. Oxidation of fuels by the electron transport chain generates a proton electrochemical gradient across the inner mitochondrial membrane. Re-entry of protons via ATP synthesis drives conversion of ADP to ATP. Uncoupling proteins (UCPs) are inner mitochondrial membrane

transporters which dissipate the proton gradient, releasing stored energy as heat (Nicholls, D.G., *et al.*, *Physiol. Rev.*, *64*:1-64 (1984); Klingenberg, M., *et al.*, *Trneds Biochem. Sci.*, *15*:108-112 (1990)). For this reason, UCPs are potentially important determinants of metabolic efficiency. UCP1, the first uncoupling protein to be identified (Lin, C.S., *et al.*, *FEBS Lett.*, *113*:299-303 (1980); Jacobsson, A., *et al.*, *J. Biol. Chem.*, *260*:16250-16254 (1985); Bouillaud, F., *et al.*, *J. Biol. Chem.*, *261*:1487-1490 (1986)), is expressed exclusively in brown adipose tissue, an important site of energy expenditure in rodents (Himms-Hagen, J., *Prog. Lipid Res.*, *28*:67-115 (1989)). However, UCP1 may be of lesser importance in humans, in whom the amount of brown adipose tissue is limited. A second uncoupling protein, referred to UCP2, was recently identified (Fleury, C., *et al.*, *Nature Genetics*, *15*:269-272 (1997)) or UCPH (Gimeno, R.E., *et al.*, *Diabetes*, *46*:900-906 (1997)). In contrast with UCP1, UCP2 is expressed in many tissues, including sites not thought to mediate energy expenditure which occurs in response to environmental temperature or diet (adaptive thermogenesis).

A greater understanding of the genes involved in metabolism will provide new approaches and targets for regulating energy expenditure in mammals.

SUMMARY OF THE INVENTION

The present invention relates to an uncoupling protein (UCP3) gene which is selectively expressed in skeletal muscle and brown fat, two tissues involved in energy expenditure in mammals. In addition, the invention relates to an alternative form of UCP3 designated UCP3-short form (UCP3sh), which is also expressed in skeletal muscle. Skeletal muscle particularly has a capacity for energy expenditure, or adaptive thermogenesis, in humans.

As used herein, "UCP3" refers to UCP3 and UCP3sh. In particular, the present invention relates to isolated (e.g., purified, essentially pure) nucleic acids (oligonucleotides, nucleotide sequences) which encode a mammalian (e.g., human) UCP3 protein, and include for example, nucleic acids (DNA, RNA) obtained from natural sources, recombinantly produced or chemically synthesized. The nucleic acids of the present invention include nucleic acids encoding human UCP3 (SEQ ID NO: 1), human UCP3sh (SEQ ID NO: 2), mouse UCP3 (SEQ ID NO: 7) and characteristic portions thereof (e.g., probes, primers). The invention also includes complementary sequences (i.e., a complement) of SEQ ID NO: 1, SEQ ID NO: 2, SEQ ID NO: 7 and characteristic portions thereof. The nucleic acids of the present invention encompass nucleic acids encoding a human UCP3 amino acid sequence (SEQ ID NO: 3), a human UCP3sh form amino acid sequence (SEQ ID NO: 4), a mouse UCP3 amino acid sequence (SEQ ID NO: 8) and characteristic portions thereof.

The present invention further relates to isolated, recombinantly produced or synthetic nucleic acids which hybridize to the nucleic acids described herein (e.g., SEQ ID NO: 1, SEQ ID NO: 2, SEQ ID NO: 7 or characteristic portions thereof) and encode UCP3 protein (a protein having the same amino acid sequence as the amino acid sequences included herein and/or a protein which exhibits the same characteristics as the UCP3 protein described herein). In particular, the invention relates to nucleic acids which hybridize, under moderate or high stringency conditions, to SEQ ID NO: 1, SEQ ID NO: 2, SEQ ID NO: 7, characteristic portions thereof or other sequences which encode UCP3.

Also encompassed by the present invention is a nucleic acid construct comprising nucleic acid which encodes a UCP3 protein (e.g., SEQ ID NO: 1, SEQ ID NO: 2, SEQ ID NO: 7 and characteristic portions thereof), wherein the nucleic acid

of the construct is expressed when the construct is present in an appropriate host cell. In one embodiment, the nucleic acid construct of the present invention is operably linked to exogenous regulatory sequence(s) such as a promoter and/or enhancer, whereby mammalian UCP3 is expressed when the host cell is maintained under conditions suitable for expression. The present invention also relates to a host cell comprising nucleic acid encoding mammalian UCP3 protein.

Also encompassed by the present invention is a method for producing a mammalian UCP3 protein (human). In the method, a nucleic acid construct comprising a nucleotide sequence (DNA, RNA) which encodes a mammalian UCP3 protein is introduced into a host cell, resulting in production of a recombinant host cell which contains a UCP3 coding sequence operably linked to an (i.e., at least one) expression control sequence. The host cells produced are maintained in a suitable medium under conditions appropriate for the nucleotide sequence to be expressed, whereby the encoded UCP3 is produced.

The present invention also relates to isolated (e.g., purified, essentially pure) UCP3 protein and includes, for example, UCP3 protein obtained from natural sources, recombinantly produced or chemically synthesized. For example, the UCP3 protein can be human UCP3 protein (SEQ ID NO: 3), human UCP3sh (SEQ ID NO:4), mouse UCP3 protein (SEQ ID NO: 8) or functional portions thereof.

The present invention also pertains to a method of identifying agents which modulate or alter (e.g., inhibit or enhance) UCP3 activity. An inhibitor of UCP3 interferes (partially or completely) with the function or bioactivity of UCP3, directly or indirectly. An enhancer (activator) of UCP3 increases or enhances the function or bioactivity of UCP3, directly or indirectly.

In one embodiment, the present invention relates to a method of identifying an agent which alters UCP3 activity, wherein a nucleic acid construct comprising nucleic acid which encodes a mammalian UCP3 is introduced into a host cell(s). The host cells produced are maintained under conditions appropriate for expression of the encoded mammalian UCP3, whereby the nucleic acid is expressed. The host cells are then contacted with a compound to be assessed (an agent) and the mitochondrial electrical potential of the cells is detected in the presence of the compound to be assessed. Detection of a change in mitochondrial electrical potential in the presence of the agent indicates that the agent alters UCP3 activity. In a particular embodiment, the invention relates to a method of identifying an agent which is an activator of UCP3 activity wherein a nucleic acid construct comprising nucleic acid which encodes a mammalian UCP3 is introduced into a host cell(s). The host cells produced are maintained under conditions appropriate for expression of the encoded mammalian UCP3, whereby the nucleic acid is expressed. The host cells are then contacted with a compound to be assessed (an agent) and the mitochondrial electrical potential of the cells is detected in the presence of the compound to be assessed. Detection of a decrease or reduction of mitochondrial electrical potential in the presence of the agent indicates that the agent activates UCP3 activity. In another embodiment, the invention relates to a method of identifying an agent which is an inhibitor of UCP3 activity, wherein a nucleic acid construct comprising nucleic acid which encodes a mammalian UCP3 is introduced into a host cell(s). The host cells produced are maintained under conditions appropriate for expression of the encoded mammalian UCP3, whereby the nucleic acid is expressed. The host cells are then contacted with a compound to be assessed (an agent) and the

mitochondrial electrical potential of the cells is detected in the presence of the compound to be assessed. Detection of an increase of mitochondrial electrical potential in the presence of the agent indicates that the agent inhibits UCP3 activity. Methods of identifying agents which alter UCP3 activity can also be performed, as described herein, using a mixture of a membrane fraction, mitochondria and UCP3 (Jezek, *et al., J. Biol. Chem. 271*:6199-6205 (1996)).

Also encompassed by the present invention is an agent which interacts with UCP3 directly or indirectly, and inhibits or enhances UCP3 function. In one embodiment, the agent is an inhibitor which interferes with UCP3 directly (e.g., by binding UCP3) or indirectly (e.g., by blocking the ability of UCP3 to regulate thermogenesis in skeletal muscle and/or brown adipose tissue). In a particular embodiment, an inhibitor of the UCP3 protein is an antibody specific for UCP3 protein or a portion of a UCP3 protein; that is, the antibody binds the UCP3 protein. For example, the antibody can be specific for the human UCP3 protein (SEQ ID NO: 3, SEQ ID NO: 4), the mouse UCP3 protein (SEQ ID NO: 8) or functional portions thereof. Alternatively, the inhibitor can be an agent other than an antibody (e.g., small organic molecule, protein, peptide) which binds UCP3 and blocks its activity. Furthermore, the inhibitor can be an agent which mimics UCP3 structurally but lacks its function. Alternatively, the inhibitor of UCP3 can be an agent which binds to or interacts with a molecule which UCP3 normally binds with or interacts with, thus blocking UCP3 from doing so and preventing it from exerting the effects it would normally exert. In another embodiment, the agent is an enhancer of UCP3 which increases the activity of UCP3 (increases thermogenesis in skeletal muscle and/or brown adipose tissue), increases the length of time it is effective (by preventing its degradation or

otherwise prolonging the time during which it is active) or both, either directly or indirectly.

The present invention also relates to antibodies (monoclonal or polyclonal) or functional portions thereof (e.g., an antigen binding portion such as an Fv, Fab, Fab', or $F(ab')_2$ fragment) which bind mammalian UCP3.

Isolation of UCP3 makes it possible to detect UCP3 in a sample (e.g., test sample). The present invention also relates to a method of detecting mammalian UCP3 in a sample (e.g., skeletal muscle, brown adipose tissue) obtained from an individual, such as a human. In one embodiment, the sample is treated to render nucleic acids in the sample available for hybridization to a nucleic acid probe (e.g., SEQ ID NO: 1, SEQ ID NO: 2, SEQ ID NO: 7 and/or characteristic portions thereof which bind to characteristic regions of UCP3-encoding nucleic acids). The treated sample is combined with a nucleic acid probe (labeled or unlabeled) comprising or complementary to all or a characteristic portion of the nucleotide sequence encoding UCP3 protein, under conditions appropriate for hybridization of complementary nucleic acids to occur. Hybridization of nucleic acids in the treated sample with the nucleic acid probe is detected; the occurrence of hybridization indicates the presence of UCP3 protein in the sample. In another embodiment, the sample is contacted with an antibody which binds to UCP3 protein (e.g., SEQ ID NO: 3, SEQ ID NO: 4, SEQ ID NO: 8 or functional portions thereof) under conditions suitable for binding of the antibody to the mammalian UCP3. Binding of the antibody to a component of the sample is detected; binding of the antibody to a component of the sample indicates the presence of UCP3 protein in the sample.

Isolation of UCP3 also makes it possible to identify a promoter(s) and/or enhancer(s) of the UCP3 gene.

Identification of promoters and/or enhancers of the UCP3 gene allow for identification of regulators of UCP3 transcription.

In addition, the present invention relates to transgenic non human animals (e.g., mice) which lack the UCP3 gene or contain a nonfunctional UCP3 gene such that UCP3 activity is lacking (e.g., UCP3 knockout mouse). The invention also relates to methods of producing UCP3 gene knockout animals, such as mice. UCP3 knockout mice can be used to further study the UCP3 gene and to assay for inhibitors and enhancers of UCP3.

The present invention also relates to a method of inhibiting (partially, completely) protein catabolism in a mammal (e.g., human) comprising administering to the mammal an effective amount of an inhibitor of UCP3. The invention also relates to a method of enhancing protein catabolism in a mammal comprising administering to the mammal an effective amount of an enhancer of UCP3. Also encompassed by the present invention is a method of inhibiting muscle wasting in a mammal comprising administering an effective amount of an inhibitor of UCP3 to the mammal.

Discovery of the UCP3 gene provides for selective modulation (enhancement, inhibition) of the expression and/or function of the UCP3 gene in skeletal muscle and brown fat, two tissues involved in adaptive thermogenesis.

BRIEF DESCRIPTION OF THE FIGURES

Figures 1A-1C are the nucleotide sequence of human UCP3 (SEQ ID NO: 1) and three different amino acid sequences (SEQ ID NO: 27, SEQ ID NO: 28 and SEQ ID NO: 29) translated from SEQ ID NO: 1.

Figures 2A-2B are the nucleotide sequence of the UCP3-short form (UCP3sh) gene (SEQ ID NO: 2) and three different amino acid sequences (SEQ ID NO: 30, SEQ ID NO: 31 and SEQ ID NO: 32) translated from SEQ ID NO: 2.

Figure 3 is a comparison of the human UCP3 amino acid sequence (SEQ ID NO: 3), the human UCP3sh amino acid sequence (SEQ ID NO: 4), the human UCP1 amino acid sequence (SEQ ID NO: 5) and the human UCP2 amino acid sequence (SEQ ID NO: 6); sequence alignments were performed using the ALIGN program (Myers, E.W., and Miller, W., *Computer Appl. Biosci.* 4:11-17 (1988); and the Genbank accession numbers for hUCP1, hUCP2 and hUCP3 are U28480, U94592 and AF001787, respectively.

Figure 4 is a graph of the hydrophilicity plots of human UCP2 and human UCP3 showing the hydrophobicity of protein across linear sequence; hydrophilicity plots for hUCP2 and hUCP3 were generated using the methods of Kyte and Doolittle (Kyte, J. and Doolittle, R.F., *J. Mol. Biol.* 157:105-132 (1982)).

Figures 5A-5C are the nucleotide sequence of mouse UCP3 (SEQ ID NO: 7) and three different amino acid sequences (SEQ ID NO: 33, SEQ ID NO: 34, SEQ ID NO: 35) translated from SEQ ID NO: 7.

Figure 6 is the amino acid sequence of mouse UCP3 (SEQ ID NO: 8).

Figure 7 is a comparison of the mouse UCP3 amino acid sequence (SEQ ID NO: 8) with the mouse UCP1 amino acid sequence (SEQ ID NO: 9), the mouse UCP2 amino acid sequence (SEQ ID NO: 10) and the human UCP3 amino acid sequence (SEQ ID NO: 3); the attached sequence and amino acid alignments, mUCP3 is 46% identical to mUCP1, 62% identical to mUCP2 but is 82% identical to hUCP3.

Figure 8 is a graphic representation of the genomic organization of the human UCP3 gene, and shows the splice donor sequence (SEQ ID NO: 11) and splice acceptor sequence (SEQ ID NO: 12) between exons 1 and 2, the splice donor sequence (SEQ ID NO: 13) and splice acceptor sequence (SEQ ID NO: 14) between exons 2 and 3, the splice donor sequence (SEQ ID NO: 15) and splice acceptor sequence (SEQ ID NO:

16) between exons 3 and 4, the splice donor sequence (SEQ ID NO: 17) and splice acceptor sequence (SEQ ID NO: 18) between exons 4 and 5, the splice donor sequence (SEQ ID NO: 19) and splice acceptor sequence (SEQ ID NO: 20) between exons 5 and 6, and the splice donor sequence (SEQ ID NO: 21) and splice acceptor sequence (SEQ ID NO: 22) between exons 6 and 7 of the UCP3 gene.

DETAILED DESCRIPTION OF THE INVENTION

The present invention relates to an uncoupling protein (UCP3) gene which is selectively expressed in skeletal muscle and brown fat, two tissues involved in energy expenditure in mammals. In addition, the invention relates to an alternative form of UCP3 designated UCP3-short form (UCP3sh), which is also expressed in skeletal muscle. As used herein, "UCP3" refers to UCP3 and UCP3sh.

The present invention relates to isolated (e.g., purified, essentially pure) UCP3 gene which is involved in regulation of thermogenesis (energy expenditure) in mammals. In particular, the present invention relates to nucleic acids (e.g., DNA, RNA, oligonucleotides, polynucleotides) or characteristic portions thereof as described herein, obtained from natural sources, recombinantly produced or chemically synthesized which encode a mammalian UCP3 or functional portion thereof.

Nucleic acids referred to herein as "isolated" are nucleic acids substantially free of (separated away from) the nucleic acids of the genomic DNA or cellular RNA of their biological source of origin (e.g., as it exists in cells or in a mixture of nucleic acids such as a library), and may have undergone further processing. "Isolated" nucleic acids include nucleic acids obtained by methods described herein, similar methods or other suitable methods, including essentially pure nucleic acids, nucleic acids produced by chemical synthesis or by combinations of

biological and chemical methods, and recombinantly produced nucleic acids which are isolated (see e.g., Daugherty, B.L. *et al., Nucleic Acids Res., 19(9)*:2471-2476 (1991); Lewis, A.P. and J.S. Crowe, *Gene, 101*: 297-302 (1991)). Nucleic acids referred to herein as "recombinant" are nucleic acids which have been produced by recombinant DNA methodologies (recombinantly produced). Recombinant DNA methodologies include, for example, expression of UCP3 in a host cell containing or modified to contain DNA or RNA encoding UCP3 or expression of UCP3 using polymerase chain reaction (PCR) techniques.

This invention includes characteristic portions of the nucleic acids described herein. As used herein, a "characteristic portion" of nucleic acids described herein refers to portions of a nucleotide sequence which encode a protein or polypeptide having at least one property, function or activity characteristic of UCP3 protein (e.g., predominantly expressed in brown adipose tissue and skeletal muscle; activity in regulating thermogenesis in skeletal muscle and brown adipose tissue; selectively uncoupling mitochondrial respiration in brown adipocytes and skeletal muscle). In addition, the term includes a nucleotide sequence which, through the degeneracy of the genetic code, encodes the same peptide as a peptide whose sequence is presented herein (e.g., SEQ ID NO: 1, SEQ ID NO: 2, SEQ ID NO: 7). The nucleic acids described herein may also contain a modification of the molecule such that the resulting gene product is sufficiently similar to that encoded by the unmodified sequence that it has essentially the same activity as the unmodified sequence. An example of such a modification would be a "silent" codon substitution or an amino acid substitution, for instance, substitution of one codon encoding a hydrophobic amino acid to another codon encoding the same hydrophobic amino acid or substitution of one acidic amino acid for another acidic

amino acid. See Ausubel, F.M., *et al.*, *Current Protocols in Molecular Biology*, Greene Publ. Assoc. and Wiley-Interscience 1989.

In one embodiment, the nucleic acid or characteristic portion thereof encodes a protein or polypeptide having at least one property, activity or function characteristic of a mammalian UCP3 (as defined herein), such as activity or function characteristic of a mammalian UCP3 (as defined herein), such as activity in regulation of thermogenesis in skeletal muscle and brown adipose tissue.

The present invention also relates more specifically to isolated nucleic acids or a characteristic portion thereof, which encode mammalian UCP3 or variants thereof.

The invention relates to isolated nucleic acids that:

(1) hybridize to (a) a nucleic acid encoding a mammalian UCP3 (e.g., human), such as a nucleic acid having a nucleotide sequence as set forth or substantially as set forth in Figures 1A-1C (SEQ ID NO:1), Figures 2A-2B (SEQ ID NO: 2) or Figures 5A-5C (SEQ ID NO: 7); (b) the complement of the sequences of (a); or (c) characteristic portions of either of the foregoing (e.g., a portion comprising the open reading frame);

(2) encode a protein or polypeptide having at least one property, activity of function characteristic of a UCP3 protein (e.g., predominantly expressed in brown adipose tissue and skeletal muscle; activity in regulating thermogenesis in skeletal muscle and brown adipose tissue; selectively uncoupling mitochondrial respiration in brown adipocytes and skeletal muscle)

(3) encode a polypeptide having the amino acid sequence of a mammalian UCP3 (e.g., SEQ ID NO: 3, SEQ ID NO: 4, SEQ ID NO: 7); or

(4) have a combination of these characteristics.

In one embodiment, the nucleic acid shares at least about 75% nucleotide sequence similarity, and more

preferably, at least about 90% nucleotide sequence similarity, to the sequence shown in Figures 1A-1C (SEQ ID NO:1), Figures 2A-2B (SEQ ID NO: 2) or Figures 5A-5C (SEQ ID NO: 7).

Isolated nucleic acids meeting these criteria include nucleic acids having sequences identical to sequences of naturally occurring mammalian UCP3 or variants of the naturally occurring sequences which encode mammalian (human) UCP3. Such variants include mutants differing by the addition, deletion or substitution of one or more residues, modified nucleic acids in which one or more residues are modified (e.g., DNA or RNA analogs), and mutants comprising one or more modified residues.

Nucleic acids of the present invention may be RNA or DNA (e.g., cDNA, genomic DNA, and synthetic DNA). The DNA may be double-stranded or single-stranded and, if single stranded, may be the coding strand or non-coding (anti-sense) strand. The coding sequence which encodes the polypeptide may be identical to the coding sequence shown in Figures 1A-1C (SEQ ID NO:1), Figures 2A-2C (SEQ ID NO:2), Figures 5A-5C (SEQ ID NO: 7) or may be a different coding sequence which, as a result of the redundancy or degeneracy of the genetic code, encodes the same polypeptide as the polypeptide encoded by the DNA of Figures 1A-1C (SEQ ID NO:1), Figures 2A-2B (SEQ ID NO:2) or Figures 5A-5C (SEQ ID NO: 7).

The nucleic acid (polynucleotide) which encodes a UCP3 polypeptide encoded by the UCP3 cDNA may include: only the coding sequence of a polypeptide; the coding sequence for a polypeptide and additional coding sequence such as a leader or secretory sequence; the coding sequence for a polypeptide (and optionally additional coding sequence) and non-coding sequence, such as introns or non-coding sequence 5′ and/or 3′ of the coding sequence.

Nucleic acids of the present invention, including those which hybridize to a selected nucleic acid as described above, can be detected or isolated under high stringency conditions or moderate stringency conditions, for example. "High stringency conditions" and "moderate stringency conditions" for nucleic acid hybridizations are explained at pages 2.10.1-2.10.16 (see particularly 2.10.8-11) and pages 6.3.1-6 in *Current Protocols in Molecular Biology* (Ausubel, F.M. *et al.*, eds., Vol. 1, Suppl. 26, 1991), the teachings of which are hereby incorporated by reference. Factors such as probe length, base composition, percent mismatch between the hybridizing sequences, temperature and ionic strength influence the stability of nucleic acid hybrids. Thus, high or moderate stringency conditions can be determined empirically, and depend in part upon the characteristics of the known nucleic acid (e.g., DNA) and the other nucleic acids to be assessed for hybridization thereto.

Nucleic acids of the present invention that are characterized by their ability to hybridize (e.g., under high or moderate stringency conditions) to (a) a nucleic acid encoding a mammalian UCP3 (for example, the nucleic acid depicted in Figures 1A-1C (SEQ ID NO:1), Figures 2A-2B (SEQ ID NO:2), Figures 5A-5B (SEQ ID NO: 7) or characteristic portions thereof); (b) the complement of the nucleic acids of (a); or (c) a portion thereof, can also encode a protein or polypeptide having at least one property, activity or function characteristic of a mammalian UCP3 as defined herein, such as activity in regulation of thermogenesis in skeletal muscle and brown adipose tissue. In a preferred embodiment the nucleic acid encodes a polypeptide which retains substantially the same biological function or activity as the polypeptide encoded by the DNA of Figures 1A-1C (SEQ ID NO:1), or Figures 2A-2B (SEQ ID NO:2) or Figures 5A-5C (SEQ ID NO: 7).

Nucleic acids of the present invention can be used in the production of proteins or polypeptides. For example, a nucleic acid (e.g., DNA) encoding a mammalian UCP3 can be incorporated into various constructs and vectors created for further manipulation of sequences or for production of the encoded polypeptide in suitable host cells as described above.

A further embodiment of the invention is antisense nucleic acid, which is complementary, in whole or in part, to a UCP3 sense strand, and can hybridize with it. The antisense strand hybridizes to DNA, or its RNA counterpart (i.e., wherein T residues of the DNA are U residues in the RNA counterpart). When introduced into a cell, antisense nucleic acid hybridizes to and inhibits the expression of the sense strand. Antisense nucleic acids can be produced by standard techniques.

In another embodiment, the antisense nucleic acid is wholly or partially complementary to and can hybridize with a target nucleic acid which encodes a mammalian UCP3. For example, antisense nucleic acid can be complementary to a target nucleic acid having the sequence shown as the open reading frame in Figures 1A-1C (SEQ ID NO:1), Figures 2A-2B (SEQ ID NO:2), Figures 5A-5C (SEQ ID NO: 7) or to a portion thereof sufficient to allow hybridization.

The nucleic acids can also be used as probes (e.g., for *in situ* hybridization) to assess regulation of thermogenesis in skeletal muscle and/or brown adipose tissue. The nucleic acids can also be used as probes to detect and/or isolate (e.g., by hybridization with RNA or DNA) polymorphic or allelic variants, for example, in a sample (e.g., skeletal muscle, brown adipocytes, white blood cells) obtained from a host (e.g., a human). Moreover, the presence or level of a particular variant in a sample(s) obtained from an individual, as compared with the presence or level in a sample(s) from normal

individuals, can be indicative of an association between abnormal regulation of thermogenesis (e.g., obesity) and a particular variant, which in turn can be used in the diagnosis of the condition.

The present invention also relates to isolated (e.g., pure, essentially pure) proteins or polypeptides designated mammalian UCP3 and variants of mammalian UCP3. In a preferred embodiment, the isolated proteins of the present invention have at least one property, activity or function characteristic of a mammalian UCP3 (as defined herein), such as activity in regulating (mediating) thermogenesis in skeletal muscle and brown adipose tissue or selectively uncoupling mitochondrial respiration in brown adipocytes and in skeletal muscle.

Proteins or polypeptides referred to herein as "isolated" are proteins or polypeptides purified to a state beyond that in which they exist in mammalian cells. "Isolated" proteins or polypeptides include proteins or polypeptides obtained by methods described herein, similar methods or other suitable methods. They include essentially pure proteins or polypeptides, proteins or polypeptides produced by chemical synthesis (e.g., synthetic peptides), or by combinations of biological and chemical methods, and recombinant proteins or polypeptides which are isolated. The proteins can be obtained in an isolated state of at least about 50 % by weight, preferably at least about 75 % by weight, and more preferably, in essentially pure form. Proteins or polypeptides referred to herein as "recombinant" are proteins or polypeptides produced by the expression of recombinant nucleic acids.

As used herein, "mammalian UCP3" protein refers to naturally occurring or endogenous mammalian UCP3s, proteins having an amino acid sequence which is the same as that of a naturally occurring or endogenous corresponding mammalian UCP3 (e.g., recombinant proteins), and functional variants

of each of the foregoing (e.g., functional fragments and/or mutants produced via mutagenesis and/or recombinant techniques). Accordingly, as defined herein, the term includes mammalian UCP3, glycosylated or unglycosylated UCP3, polymorphic or allelic variants, and other isoforms of mammalian UCP3 (e.g., produced by alternative splicing or other cellular processes), and functional fragments.

Naturally occurring or endogenous mammalian UCP3s include wild type proteins such as mammalian UCP3, polymorphic or allelic variants and other isoforms which occur naturally in mammals (e.g., primate, preferably human, murine, bovine). Such proteins can be recovered from a source in which UCP3 is naturally produced. for example. These mammalian proteins have the same amino acid sequence as naturally occurring or endogenous corresponding mammalian UCP3.

"Functional variants" of mammalian UCP3 include functional fragments, functional mutant proteins, and/or functional fusion proteins. Generally, fragments or portions of mammalian UCP3 encompassed by the present invention include those having one or more amino acid deletions relative to the naturally occurring mammalian UCP3 protein (such as N-terminal, C-terminal or internal deletions). Fragments or portions in which only contiguous amino acids have been deleted or in which non-contiguous amino acids have been deleted relative to naturally occurring mammalian UCP3 are also encompassed by the invention.

Generally, mutants or derivatives of mammalian UCP3, encompassed by the present invention include natural or artificial variants differing by the addition, deletion and/or substitution of one or more contiguous or non-contiguous amino acid residues, or modified polypeptides in which one or more residues is modified, and mutants comprising one or more modified residues. For

example, mutants can be natural or artificial variants of mammalian UCP3 which differ from naturally occurring UCP3 by the addition, deletion and/or substitution of one or more contiguous or non-contiguous amino acid residues.

A "functional fragment or portion", "functional mutant" and/or "functional fusion protein" of a mammalian UCP3 refers to an isolated protein or oligopeptide which has at least one property, activity or function characteristic of a mammalian UCP3, such as activity in regulating (mediating) thermogenesis in skeletal muscle and brown adipose tissue or activity in selectively uncoupling mitochondrial respiration in brown adipocytes and in skeletal muscle.

Suitable fragments or mutants can be identified by screening. For example, the N-terminal, C-terminal, or internal regions of the protein can be deleted in a step-wise fashion and the resulting protein or polypeptide can be screened using a suitable assay, for example, by measuring mitochondrial membrane potential in a host cell expressing UCP3. Where the resulting protein displays activity in the assay, the resulting protein ("fragment") is functional.

The invention also encompasses fusion proteins, comprising a mammalian UCP3 as a first moiety, linked to a second moiety not occurring in the mammalian UCP3 found in nature. Thus, the second moiety can be, for example, an amino acid, oligopeptide or polypeptide. The first moiety can be in an N-terminal location, C-terminal location or internal location of the fusion protein. In one embodiment, the fusion protein comprises a mammalian UCP3 or portion thereof as the first moiety, and a second moiety comprising an affinity ligand (e.g., an enzyme, an antigen, epitope tag) joined to the first moiety. Optionally, the two components can be joined by a linker.

Examples of "human UCP3" include proteins having an amino acid sequence as set forth or substantially as set forth in Figure 3 (SEQ ID NO: 3, SEQ ID NO: 4) and functional portions thereof. An example of "mouse UCP3" includes a protein having an amino acid sequence as set forth or substantially set forth in Figure 6 (SEQ ID NO: 8). In preferred embodiments, a human UCP3 protein, a mouse UCP3 protein or a variant thereof has an amino acid sequence which has at least about 75% identity, and more preferably at least about 90% identity, to the protein shown in Figure 3 (SEQ ID NO: 3, SEQ ID NO: 4) or Figure 6 (SEQ ID NO: 8).

Another aspect of the invention relates to a method of producing a human UCP3 or variant (e.g., portion) thereof. Recombinant protein can be obtained, for example, by the expression of a recombinant DNA molecule encoding a mammalian UCP3 or variant thereof in a suitable host cell.

Constructs suitable for the expression of a mammalian UCP3 or variant thereof are also provided. The constructs can be introduced into a suitable host cell, and cells which express a recombinant mammalian UCP3 or variant thereof, can be produced and maintained in culture. Such cells are useful for a variety of purposes, and can be used in the production of protein for characterization, isolation and/or purification, (e.g., affinity purification), and as immunogens, for instance. Suitable host cells can be procaryotic, including bacterial cells such as *E. coli*, *B. subtilis* and or other suitable bacteria (e.g., *Streptococci*) or eucaryotic, such as fungal or yeast cells (e.g., *Pichia pastoris*, *Aspergillus species*, *Saccharomyces cerevisiae*, *Schizosaccharomyces pombe*, *Neurospora crassa*), or other lower eucaryotic cells, and cells of higher eucaryotes such as those from insects (e.g., Sf9 insect cells) or mammals (e.g., Chinese hamster ovary cells (CHO), COS cells, HuT 78 cells, 293 cells).

(See, e.g., Ausubel, F.M. *et al.*, eds. *Current Protocols in Molecular Biology*, Greene Publishing Associates and John Wiley & Sons Inc., (1993)).

Host cells which produce a recombinant mammalian UCP3 or variants thereof can be produced as follows. For example, nucleic acid encoding all or part of the UCP3 protein or a functional portion thereof can be inserted into a nucleic acid vector, e.g., a DNA vector, such as a plasmid, virus or other suitable replicon for expression. A variety of vectors is available, including vectors which are maintained in single copy or multiple copy, or which become integrated into the host cell chromosome.

The transcriptional and/or translational signals of a mammalian UCP3 gene can be used to direct expression. Alternatively, suitable expression vectors for the expression of a nucleic acid encoding all or part of the desired protein are available. Suitable expression vectors can contain a number of components, including, but not limited to, one or more of the following: an origin of replication; a selectable marker gene; one or more expression control elements, such as a transcriptional control element (e.g., a promoter, an enhancer, terminator), and/or one or more translation signals; a signal sequence or leader sequence for membrane targeting or secretion (of mammalian origin or from a heterologous mammal or non-mammalian species). In a construct, a signal sequence can be provided by the vector, the mammalian UCP3 coding sequence, or other source.

A promoter can be provided for expression in a suitable host cell. Promoters can be constitutive or inducible. The promoter is operably linked to nucleic acid encoding the mammalian UCP3 or variant thereof, and is capable of directing expression of the encoded polypeptide in the host cell. A variety of suitable promoters for procaryotic (e.g., lac, tac, T3, T7 promoters for *E. coli*)

and eucaryotic (e.g., yeast alcohol dehydrogenase (ADH1), SV40, CMV) hosts is available.

In addition, the expression vectors typically comprise a selectable marker for selection of host cells carrying the vector, and in the case of a replicable expression vector, also comprise an origin of replication. Genes encoding products which confer antibiotic or drug resistance are common selectable markers and may be used in procaryotic (e.g., β-lactamase gene (ampicillin resistance), *Tet* gene for tetracycline resistance) and eucaryotic cells (e.g., neomycin (G418 or geneticin), gpt (mycophenolic acid), ampicillin, or hygromycin resistance genes). Dihydrofolate reductase marker genes permit selection with methotrexate in a variety of hosts. Genes encoding the gene product of auxotrophic markers of the host (e.g., *LEU2, URA3, HIS3*) are often used as selectable markers in yeast. Use of viral (e.g., baculovirus) or phage vectors, and vectors which are capable of integrating into the genome of the host cell, such as retroviral vectors, are also contemplated. The present invention also relates to cells carrying these expression vectors.

For example, a nucleic acid encoding a mammalian UCP3 or variant thereof is incorporated into a vector, operably linked to one or more expression control elements, and the construct is introduced into host cells which are maintained under conditions suitable for expression, whereby the encoded polypeptide is produced. The construct can be introduced into cells by a method appropriate to the host cell selected (e.g., transformation, transfection, electroporation, infection). For production of a protein, host cells comprising the construct are maintained under conditions appropriate for expression, (e.g., in the presence of inducer, suitable media supplemented with appropriate salts, growth factors, antibiotic, nutritional

supplements, etc.). The encoded protein (e.g., human UCP3) can be isolated from the host cells or medium.

Fusion proteins can also be produced in this manner. For example, some embodiments can be produced by the insertion of a mammalian UCP3 cDNA or portion thereof into a suitable expression vector, such as Bluescript®II SK +/- (Stratagene), pGEX-4T-2 (Pharmacia), pcDNA-3 (Invitrogen) and pET-15b (Novagen). The resulting construct can then be introduced into a suitable host cell for expression. Upon expression, fusion protein can be isolated or purified from a cell lysate by means of a suitable affinity matrix (see e.g., *Current Protocols in Molecular Biology* (Ausubel, F.M. *et al.*, eds., Vol. 2, Suppl. 26, pp. 16.4.1-16.7.8 (1991)). In addition, affinity labels provide a means of detecting a fusion protein. For example, the cell surface expression or presence in a particular cell fraction of a fusion protein comprising an antigen or epitope affinity label can be detected by means of an appropriate antibody.

The UCP3 nucleic acids (DNA, RNA) and protein can be used in a variety of ways. For example, UCP3 nucleic acids and proteins can be used to identify agents (e.g., molecules) that alter or modulate (enhance, inhibit) UCP3 expression and/or function. For example, UCP3 can be expressed in a host cell and effects of test compounds on mitochondrial membrane potential in the host cell could be assessed. In addition, evaluation of mitochondrial respiration could also be performed in the host cell.

In one embodiment, the present invention relates to a method of identifying an agent which alters UCP3 activity, wherein a nucleic acid construct comprising nucleic acid which encodes a mammalian UCP3 is introduced into a host cell(s). The host cells produced are maintained under conditions appropriate for expression of the encoded mammalian UCP3, whereby the nucleic acid is expressed. The host cells are then contacted with a compound to be

assessed (an agent) and the mitochondrial electrical potential (mitochondrial membrane potential) of the cells is detected in the presence of the compound to be assessed. Detection of a change in mitochondrial electrical potential in the presence of the agent indicates that the agent alters UCP3 activity. In a particular embodiment, the invention relates to a method of identifying an agent which is an activator of UCP3 activity wherein a nucleic acid construct comprising nucleic acid which encodes a mammalian UCP3 is introduced into a host cell(s). The host cells produced are maintained under conditions appropriate for expression of the encoded mammalian UCP3, whereby the nucleic acid is expressed. The host cells are then contacted with a compound to be assessed (an agent) and the mitochondrial electrical potential of the cells is detected in the presence of the compound to be assessed. Detection of a decrease or reduction of mitochondrial electrical potential in the presence of the agent indicates that the agent activates UCP3 activity. In another embodiment, the invention relates to a method of identifying an agent which is an inhibitor of UCP3 activity, wherein a nucleic acid construct comprising nucleic acid which encodes a mammalian UCP3 is introduced into a host cell(s). The host cells produced are maintained under conditions appropriate for expression of the encoded mammalian UCP3, whereby the nucleic acid is expressed. The host cells are then contacted with a compound to be assessed (an agent) and the mitochondrial electrical potential of the cells is detected in the presence of the compound to be assessed. Detection of an increase of mitochondrial electrical potential in the presence of the agent indicates that the agent inhibits UCP3 activity.

Detection of a change in mitochondrial electrical potential can be performed using a variety of techniques. For example, a change in mitochondrial electrical potential

can be detected by measuring fluorescence of recombinant cells expressing UCP3. Decrease of fluorescence in the presence of the test compound, indicates a decrease of mitochondrial membrane potential (mitochondrial $\Delta\Psi$), and vice versa for cases where fluorescence is increased. That is, increase of fluorescence in the presence of the test compound indicates an increase of mitochondrial $\Delta\Psi$. If decrease in fluorescence is observed in UCP3 expressing cells, but not in control cells, then the test compound is an activator of UCP3. If an increase in fluorescence is observed in UCP3 expressing cells, but not in control cells, then the test compound is an inhibitor of UCP3.

In a particular embodiment, as described in Example 3, a high throughput screen can be used to identify agents that activate (enhance) or inhibit UCP3 activity. For example, the method of identifying an agent which alters UCP3 activity can be performed as follows. A nucleic acid construct comprising nucleic acid which encodes a mammalian UCP3 is introduced into a host cell(s) to produce recombinant host cells. The recombinant host cells produced are maintained under conditions appropriate for expression of the encoded mammalian UCP3, whereby the nucleic acid is expressed. A fluorescent dye and the compound to be assessed are added to the recombinant host cells; the resulting combination is referred to as a test sample. Fluorescence is detected. A decrease of fluorescence in the presence of the test compound occurs with a decrease in the mitochondrial electrical potential of the cells, which indicates that the agent is an activator of UCP3. Conversely, an increase of fluorescence in the presence of the test compound occurs with an increase in the mitochondrial electrical potential of the cells, which indicates that the agent is an inhibitor of UCP3. Suitable dyes for use in this embodiment of the

invention include, for example, JC-1, rhodamine 123, DiOCc[3], or tetramethylhydrosamine.

A control can be used in the methods of detecting agents which alter UCP3 activity. For example, the control sample includes the same reagents but lacks the compound or agent being assessed; it is treated in the same manner as the test sample.

Also encompassed by the present invention is an agent which interacts with UCP3 directly or indirectly, and inhibits or enhances UCP3 expression and/or function. In one embodiment, the agent is an inhibitor which interferes with UCP3 directly (e.g., by binding UCP3) or indirectly (e.g., by blocking the ability of UCP3 to function in thermogenesis). In a particular embodiment, an inhibitor of UCP3 protein is an antibody specific for UCP3 protein or a functional portion of UCP3; that is, the antibody binds the UCP3 protein. For example, the antibody can be specific for the protein encoded by the amino acid sequence of human UCP3 (SEQ ID NO: 3), human UCP3sh (SEQ ID NO: 4), mouse UCP3 (SEQ ID NO: 8) or portions thereof. Alternatively, the inhibitor can be an agent other than an antibody (e.g., small organic molecule, protein or peptide) which binds UCP3 and blocks its activity. For example, the inhibitor can be an agent which mimics UCP3 structurally, but lacks its function. Alternatively, it can be an agent which binds to or interacts with a molecule which UCP3 normally binds with or interacts with, thus blocking UCP3 from doing so and preventing it from exerting the effects it would normally exert.

In another embodiment, the agent is an enhancer (activator) of UCP3 which increases the activity of UCP3 (increases the effect of a given amount or level of UCP3), increases the length of time it is effective (by preventing its degradation or otherwise prolonging the time during which it is active) or both either directly or indirectly.

For example, UCP3 nucleic acids and proteins can be used to identify anti-obesity drugs which enhance UCP3 to induce uncoupling in brown fat and/or skeletal muscle, with the result that stored energy is released as heat.

In another embodiment, the sequences described herein can be used to detect UCP3 or DNA encoding UCP3 in a sample. For example, a labeled nucleic acid probe having all or a functional portion of the nucleotide sequence of UCP3 can be used in a method to detect UCP3 in a sample. In one embodiment, the sample is treated to render the nucleic acids in the sample available for hybridization to a nucleic acid probe, which can be DNA or RNA. The resulting treated sample is combined with a labeled nucleic acid probe having all or a portion of the nucleotide sequence of UCP3, under conditions appropriate for hybridization of complementary sequences to occur. Detection of hybridization of nucleic acids from the sample with the labeled nucleic probe indicates the presence of UCP3 in a sample. The presence of UCP3 mRNA is indicative of UCP3 expression. Such a method can be used, for example, as a screen for normal or abnormal thermogenesis in skeletal muscle or brown adipose tissue.

Alternatively, a method of detecting UCP3 in a sample can be accomplished using an antibody directed against UCP3 or a portion of UCP3. Detection of specific binding to the antibody indicates the presence of UCP3 in the sample (e.g., ELISA). This could reflect a pathological state associated with UCP3 and, thus, can be used diagnostically.

The sample for use in the methods of the present invention includes a suitable sample from, for example, a mammal, particularly a human. For example, the sample can be blood, skeletal muscle or brown adipose tissue.

The UCP3 sequences of the present invention can also be used to generate nonhuman gene knockout animals, such as mice, which lack UCP3 and transgenically overexpress UCP3.

For example, such UCP3 gene knockout mice can be generated and used to obtain further insight into the function of UCP3 as well as assess the specificity of UCP3 activators and inhibitors. Also, overexpression of UCP3 (e.g., human UCP3) in transgenic mice can be used as a means of creating a test system for UCP3 activators and inhibitors (e.g., against human UCP3). In addition, the UCP3 gene can be used to clone the UCP3 promoter/enhancer in order to identify regulators of UCP3 transcription. UCP3 gene knockout animals include animals which completely or partially lack the UCP3 gene and/or UCP3 activity or function.

As described herein, it is likely that UCP3 plays a role in controlling protein wasting and production of gluconeogenic precursors by skeletal muscle via transport of one or more metabolites, which indicates that inhibitors of UCP3 can be used as a means of curtailing muscle wasting due to, for example, infection, (e.g., human immunodeficiency virus) cancer, tumor cachexia, muscle diseases (e.g., muscular dystrophy) or as a possible treatment for non-insulin dependent diabetes mellitus (NIDDM).

Thus the present invention relates to a method of inhibiting (partially, completely) protein catabolism in a mammal (e.g., human) comprising administering to the mammal an effective amount of an inhibitor of UCP3. The invention also relates to a method of enhancing protein catabolism in a mammal comprising administering to the mammal an effective amount of an enhancer UCP3. Also encompassed by the present invention is a method of inhibiting muscle wasting in a mammal comprising administering an effective amount of an enhancer of UCP3 to the mammal.

A number of studies have demonstrated that brown adipose tissue plays an important role in regulating energy balance in rodents (Himms-Hagen, J., *Prog. Lipid Res.*,

28:67-115 (1989)). The tissue is highly specialized for stimulated energy expenditure with a rich vascular supply, dense sympathetic innervation, and numerous mitochondria. Importantly, brown adipocytes are further distinguished from other cell types by their expression of all three uncoupling proteins: UCP1, which is expressed exclusively in brown adipocytes, UCP2, which is expressed widely (Fleury, C., *et al.*, *Nature Genetics*, *15*:269-272 (1997); Gimeno, R.E., *et al.*, *Diabetes, in press* (1997)) and, as demonstrated herein, UCP3 which is expressed selectively and abundantly in brown adipocytes and skeletal muscle. These features make brown fat ideally suited to regulated thermogenesis.

In contrast to rodents, brown adipose tissue in large mammals is relatively limited and therefore brown fat may not be a significant regulator of human energy expenditure. A number of studies in humans have implicated skeletal muscle as an important mediator of adaptive thermogenesis in humans (Astrup, A., *et al.*, *Am. J. Physiol.*, *248*:E507-515 (1985); Astrup, A., *et al.*, *Am. J. Physiol.*, *257*:E340-345 (1989); Zurlo, F., *et al.*, *J. Clin. Invest.*, *86*:1423-1427 (1990); Simonsen, L., *et al.*, *Am. J. Physiol.*, *263*:E850-855 (1992); Spraul, M., *et al.*, *J. Clin. Invest.*, *92*:1730-1735 (1993)). Approximately 80% of the variance in resting energy expenditure between individuals can be accounted for by differences in fat-free mass (Ravussin, E., *et al.*, *Am. J. Clin. Nutr.*, *55*:242S-245S (1992)), much of which is skeletal muscle. Similarly, a perfused forearm study has demonstrated that differences in skeletal muscle energy expenditure account for much of the variation in metabolic rate observed between individuals (Zurlo, F., *et al.*, *J. Clin. Invest.*, *86*:1423-1427 (1990)). Regulated energy expenditure in skeletal muscle is controlled, in large part, by sympathetic stimulation ((Astrup, A., *et al.*, *Am. J. Physiol.*, *248*:E507-515 (1985); Astrup, A., *et*

al., *Am. J. Physiol.*, *257*:E340-345 (1989); Simonsen, L., *et al.*, *Am. J. Physiol.*, *263*:E850-855 (1992); Spraul, M., *et al.*, *J. Clin. Invest.*, *92*:1730-1735 (1993)). It is interesting to note that brown fat and skeletal muscle have many features in common: a rich blood supply, a dense sympathetic innervation, and abundant mitochondria. In addition, both tissues express high levels of UCP3 mRNA.

The heart continuously expends large amounts of energy in order to maintain blood circulation. In view of this, it is probably significant that UCP3 is minimally expressed in cardiac tissue. This is especially true given the general tendency for non-contractile muscle-specific genes to be expressed in both striated muscle types (skeletal and cardiac). Abundant expression of UCP3 in two thermogenic tissues, skeletal muscle and brown fat, and relative lack of expression in other sites such as the heart, demonstrates that UCP3 is an important molecular mediator of adaptive thermogenesis.

Thus, the present invention provides for anti-obesity drug development wherein the UCP3 nucleic acids and protein can be used to identify, for example, enhancers (activators) of UCP3 which can be used to induce uncoupling. β3-adrenergic receptor agonists, which increase UCP1 expression and activity in brown fat are presently under development, but may have limited effects given the paucity of brown fat in humans. UCP2 is another potential target. However, it is expressed in a number of critical organs and tissues and its activation could produce unwanted side effects. Specific activators of UCP3 expression and/or function, on the other hand, selectively increase energy expenditure in skeletal muscle and brown fat, two tissues that have the capacity for adaptive energy expenditure.

The present invention is further illustrated by the following examples, which are not intended to be limiting in any way.

EXAMPLE 1 CLONING AND CHARACTERIZATION OF THE UCP3 GENE

Northern Blot Assays

Human Multiple Tissue Northern Blots (#7760-1, #7759-1 and #7767-1) containing approximately 2 μg of polyA RNA per lane were purchased from Clontech Laboratories (Palo Alto, CA). All hybridizations, membranes washes and membrane strippings were performed according to manufacturer's specifications. The blots were first hybridized to a hUCP3 probe, washed and exposed to film for 1-18 hours, then stripped, rehybridized to a hUCP2 probe and exposed to film for 18 hours. The hUCP3 probe was a 293 bp fragment corresponding to residues #211-308. The hUCP2 probe was a 1125 bp fragment spanning the entire open reading frame. The specific activities of both hybridization probes were similar. Mouse Northern blots were generated using total RNA isolated from a number of tissues and equal loading of lanes was established using ethidium bromide florescence. The mouse Northern blots were hybridized using the hUCP3 probe described above.

RNase Protection

Total RNA was extracted from adipose tissue the method of Chomczynski and Sacchi (Chomczynski, P. and Sacchi, N., *Anal. Biochem.*, *162*:156-159 (1987)). Skeletal muscle and heart RNA was obtained from Clontech. Aliquots of 1, 3, 5 and 10 μg of adipose tissue and skeletal muscle RNA and 10 μg aliquot of heart RNA were used for determination of UCP3 and mRNA levels. The Rnase protection assay was performed as previously described (Vidal-Puig, A., *et al.*, *J. Clin. Invest.*, *97*:2553-2561 (1997)). A UCP-3 cDNA fragment was

generated by reverse transcriptase-PCR using total RNA from human muscle as follows: two primers (5'GGA CTA CCA CCT GCT CAC TG 3'(SEQ ID NO: 23) and 5' CCC GTA ACA TAT GGA CTT T3' (SEQ ID NO: 24)) were designed to amplify 302 bp of the hUCP-3 sequence corresponding to residues #209-308. The PCR product was subcloned into PGMT easy TA cloning vector (Promega Corp., Madison, WI) and linearized for riboprobe synthesis using *Spe* I. Identity and orientation of the UCP3 probe was confirmed by sequencing. The antisense [^{32}P]-labeled UCP3 template was synthesized using T& RNA polymerase. A human cyclin riboprobe was used as an internal control (Ambion, Inc., Austin, TX).

Results

As described herein, a third uncoupling homologue designated UCP3 has been cloned. It is distinguished from UCP1 and UCP2 by its selective expression in skeletal muscle and brown adipose tissue, two important sites for regulated energy expenditure in humans (Astrup, A., *et al.*, *Am. J. Physiol.*, *248*:E507-515 (1985); Astrup, A., *et al.*, *Am. J. Physiol.*, *257*:E340-345 (1989); Zurlo, F., *et al.*, *J. Clin. Invest.*, *86*:1423-1427 (1990); Simonsen, L., *et al.*, *Am. J. Physiol.*, *263*:E850-855 (1992); Spraul, M., *et al.*, *J. Clin. Invest.*, *92*:1730-1735 (1993)) and rodents (Himms-Hagen, J., *Prog. Lipid Res.*, *28*:67-115 (1989)). At the amino acid level, hUCP3 is 71% identical to hUCP2 and 57% identical to hUCP1. Because UCP3 is abundantly and selectively expressed in skeletal muscle and brown adipose tissue, UCP3 is likely to be an important mediator of regulated thermogenesis in humans. Since UCP3 is minimally expressed in heart and other critical organs, it is a promising target for anti-obesity drug development aimed at increasing thermogenesis.

The expressed sequence tag (EST) database (http://www.ncbi.nlm.gov) was screened for sequences homologous to UCP1. One human EST, deposited by the Washington University, St. Louis - Merck & Co. EST project, was identified which was similar but not identical to hUCP1 and hUCP2 (accession no. AA192136, IMAGE clone no. 628529). This clone originated from a human skeletal muscle cDNA library (#937209, Stratagene, La Jolla, CA). The bacterial stock for clone 628529 was obtained from Genome Systems (St. Louis, MI) and was found to contain an insert of approximately 1.3kb, which included the C-terminal third of the open reading frame. The coding region within clone 628529 was fully resequenced. Full-length cDNA sequences were generated using the Marathon cDNA Amplification Kit, human skeletal muscle Marathon-Ready cDNA (both from Clontech Laboratories, Palo Alto, CA) and an antisense primer (5'-TTC ACC ACG TCC ACC CGG GGG GAT GCC ACC-3') (SEQ ID NO: 25) corresponding to the coding sequence presumed to represent hUCP3.

UCP3 cDNA sequence contains a 5' untranslated region of at least 183 bases, an open reading from of 936 bases, a 3' untranslated region of approximately 1.1 kb, a polyadenylation signal and a polyA tail (Figures 1A-1C). The UCP3 mRNA transcript is predicted to be equal to or greater than 2.2 kb. UCP3 protein, as deduced from the open reading frame, is composed of 312 amino acids and is estimated to have a molecular weight of approximately 34 kD (Figure 3). As shown in Figure 3, at the amino acid level, hUCP3 is 71% identical to hUCP2 and 57% identical to hUCP1; and hUCP2 is 59% identical to hUCP1. Many of the nonidentical residues in hUCP3 are conservative substitutions which in most cases correspond to residues found in either mUCP2 (Fleury, C., *et al.*, *Nature Genetics*, *15*:269-272 (1997); Gimeno, R.E., *et al.*, *Diabetes*, *46*:900-906 (1997)) or in UCP1 from various species (Klaus, S., *et*

al., Int. J. Biochem., 23:791-801 (1991)). The data, based upon the high degree of homology between UCP1, UCP2 and UCP3, demonstrates that UCP3 uncouples mitochondrial respiration.

In order to establish the tissue distribution of UCP3 in humans, Northern blot analyses were performed. UCP3 was abundantly expressed in skeletal muscle, generating a dominant mRNA transcript of approximately 2.4 kb. With longer exposure (18 hours), a much weaker UCP3 signal (2.4 kb) was detected in a large number of other tissues and organs. The longer exposures (18 hours) of the human UCP3 Northern blots also revealed the presence of a smaller mRNA transcript which had a similar size (approximately 1.6 kb). Of note, the 294 bp hUCP3 probe employed was 75% identical to hUCP2. Rehybridization of the same blots with hUCP2 confirmed that this smaller 1.6 kb signal was UCP2. The UCP2 signal, as previously reported (Fleury, C., *et al., Nature Genetics, 15*:269-272 (1997); Gimeno, R.E., *et al., Diabetes, 446*:900-906 (1997)) was widely expressed. It was being most abundant in spleen, thymus, bone marrow, trachea, and lymph node, and somewhat less abundant in skeletal muscle as well as a number of other tissues. UCP2 was also abundantly expressed in white adipose tissue as reported Gimeno, R.E., *et al., Diabetes, 446*:900-906 (1997)). A comparison of hybridization signals for UCP2 and UCP3 suggests that UCP3 may be the dominant uncoupling protein transcript in human skeletal muscle.

A sensitive RNase protection assay was used to assess UCP3 mRNA expression in heart, skeletal muscle and white adipose tissue. No UCP3 signal could be detected in white adipose tissue. In heart, a very weak UCP3 signal was detected. The signal in heart was less than 1% of that detected in skeletal muscle.

In mice, abundant UCP3 expression was detected in skeletal muscle and brown fat. As with humans, little or

no UCP3 expression was detected in other mouse tissues such as white adipose tissue, brain, kidney, liver and colon. As was observed in the human mRNA studies, a smaller transcript was detected in mouse samples as well. This smaller transcript most likely represents mUCP2 given that it was most abundant in white adipose tissue, a site of high-level UCP2 expression (Fleury, C., *et al.*, *Nature Genetics, 15*:269-272 (1997); Gimeno, R.E., *et al.*, *Diabetes, in press* (1997)). Of note, the hUCP3 probe is 73% identical to mUCP2.

Figure 4 is a hydrophilicity plot of human UCP2 and human UCP3 showing the hydrophobicity of protein across linear sequence.

EXAMPLE 2 Discovery of an Alternative Form of UCP3, Designated UCP3-short form (UCP3sh)

As discussed above, the genomic organization of the human UCP3 gene has been defined. In addition, it has been determined that the UCP3 gene generates two mRNA transcripts, UCP3 and UCP3-short form (UCP3sh). The nucleotide sequence of UCP3sh mRNA is shown in Figures 2A-2B. The UCP3sh transcript encodes a shortened version of the UCP3 protein. As shown in Figure 8, the UCP3sh transcript results when a polyadenylation/transcription termination signal (AATAAA) (SEQ ID NO: 26) located within intron 6 terminates transcription (see Figure 3). However, this AATAAA (SEQ ID NO: 26) seems to be only partially effective in terminating transcription. When it does succeed in terminating transcription, the UCP3sh transcript is generated. When it fails to terminate transcription, transcription continues on through exon 7 and terminates at the exon 7 AATAAA (SEQ ID NO: 26) signal. Splicing between exon 6 and exon 7 then occurs to generate the UCP3 transcript.

As shown is Figure 3, UCP3sh differs from UCP3 only by the absence of the last 37 amino acids. It is reasonable to expect that this is significant, since the region missing in UCP3sh is highly homologous to a region in UCP1 which has been implicated in mediating inhibition of uncoupling activity by purine nucleotides (Murdza-Inglis, D.L., *et al.*, *J Biol Chem. 269*:7435-7438 (1994)). As a result, it is reasonable to expect that UCP3sh is more active as an uncoupler than UCP3.

Using a quantitative RNase protection assay similar to that described in Example 1, it was determined that UCP3sh mRNA, like UCP3 mRNA is extremely abundant in human skeletal muscle. In normal individuals, the level of UCP3sh mRNA is skeletal muscle is equal to or greater that the level of UCP3 mRNA. Preliminary studies have indicated that UCP3sh mRNA levels are reduced in obese individuals compared to lean individual. In contrast, UCP3 mRNA levels seem to be unchanged in obese individuals. These preliminary findings raise the possibility that UCP3sh is the more important UCP3 protein for body weight regulation.

EXAMPLE 3 Cloning of mouse UCP3 gene

Using the human UCP3 gene, the mouse UCP3 gene was isolated using methods similar to those described in Example 1. The mouse UCP3 nucleotide sequence (SEQ ID NO: 7) is shown in Figures 5A-5C, and the mouse UPC3 amino acid sequence is shown in Figure 6. Comparisions of mUCP3 versus mUCP1 and mUP2 and human UCP3 are shown in Figure 7.

EXAMPLE 4 Monitoring of JC-1 fluorescence in living cells

An assay which utilizes fibroblast-like cells lines expressing recombinant human UCP3, and a fluorescent dye (e.g., JC-1) makes it possible to rapidly assess

mitochondrial membrane potential ($\Delta\Psi$) in living cells (Smiley, S.T., *et al., Proc. Natl. Acad. Sci. USA, 88*:3671-3675 (1991); Reers, M., *et al., Methods in Enzymology, 260*:406-417 (1995)). Any drug which increases UCP3 activity is expected to reduce $\Delta\Psi$, and therefore, reduce "red"-fluorescence of JC-1. By comparing effects of test compounds on fluorescence in a cell line expressing UCP3 with a control (e.g., cells which do not express UCP3; cells which express UCP3 in the absence of the test compound), it is possible to identify specific activators and inhibitors of UCP3. The cells can be grown in 96 well plates, and the plates can be read directly in a fluorometer designed to handle 96 well plates, it is possible to perform this assay in a high-throughput fashion.

Recombinant cells expressing hUCP3 and cells not expressing UCP3 are grown in 96 well plates. On the day of analysis, the plates are rinsed and JC-1 dye is added to all wells plus or minus test compounds. Later, plates are washed and then, in the presence of the test compound, fluorescence is determined in a fluorometer. Decrease of fluorescence in the presence of the test compound, indicates a decrease of mitochondrial $\Delta\Psi$ (and vice versa for cases where fluorescence is increased). That is, increase of fluorescence in the presence of the test compound indicates an increase of mitochondrial $\Delta\Psi$. If decrease in fluorescence is observed in UCP3 expressing cells but not in control cells, then the test compound is an activator of UCP3. If an increase in fluorescence is observed in UCP3 expressing cells, but not in control cells, then the test compound is an inhibitor of UCP3.

Any dye can be used in the high-throughout screen, such as JC-1, rhodamine 123, DiOCc[3], or tetramethylhydrosamine. In a particular embodiment, JC-1 dye, a delocalized lipophilic cation (DLC), can be used.

The distinguishing feature of DLCs is that they are positively charged, yet lipophilic. The lipophilic feature allows then to traverse membranes and the positive charge causes then to accumulate within mitochondria (negatively charged on the inside). This accumulation is proportional to $\Delta\Psi$, the membrane electrical potential across the inner mitochondrial membrane, and follows the Nernst Equation shown below. The mitochondrial $\Delta\Psi$ results from the protein electrochemical gradient across the inner mitochondrial membrane and represents the electrical portion of this gradient (ΔpH represents the chemical portion of the gradient).

$$\Delta\Psi = -60 \log F_{in}/F_{out} \quad F = \text{concentration of DLC}$$

Thus, a $\Delta\Psi$ of -60 mV corresponds to a DLC in/out ratio of 10 to 1, and a $\Delta\Psi$ of -120 mV, corresponds to a DLC in/out ratio of 100 to 1. Thus, a change in $\Delta\Psi$ is amplified by a change in F_{in}/F_{out}. Of note, $\Delta\Psi$ for most mitochondrial range between -50 mV and -160 mV.

Protonophore uncouplers such as DNP (dinitrophenol), CCCP (carbonyl cyanide *m*-chlorophenyllhydrazone), decrease $\Delta\Psi$ and, as a result, markedly decrease the accumulation of JC-1. Any drug which increases UCP activity is expected to have the same effect as DNP, CCCP or FCCP.

JC-1 has fluorescent features which makes it extremely useful as a monitor of mitochondrial $\Delta\Psi$. Many dyes aggregate at high concentrations and this reduces fluorescence greatly (for example, rhodamine 123). Aggregates of JC-1 fluoresce intensely, and at higher wavelength than JC-1 monomers. Specifically, monomers emit at 527 nM (green) while J-aggregates emit at 590 nM (red). Thus, high concentrations of JC-1 accumulate in mitochondria permitting the formation of aggregates. The

accumulation of JC-1 and therefore the formation of aggregates is proportional to mitochondrial $\Delta\Psi$. Aggregates do not form in other cellular locations due to insufficient accumulation of JC-1. Thus, detection of aggregates (as measured by fluorescence at 590 nM) is a sensitive indicator of mitochondrial $\Delta\Psi$.

CX-1 cells were incubated with JC-1 (10ug/ml) with or without the uncoupler, FCCP, for 10 minutes, washed 3 times, trypsinized and then transferred as a cell suspension to a 1 cm quartz cuvette, in which fluorescence was monitored using a Kontron SFM25 fluorescent spectrophotometer.

	Fluorescence (in arbitrary units) 520 nM (green)	590 nM (red)
CX-1 cells	90	90
CX-1 cells + FCCP	80	10

The data shows that JC-1 aggregate fluorescence can be monitored in living cells and that an uncoupler (FCCP) which is expected to have the same effect as a UCP activator markedly lowers "red" fluorescence. Fluorescence can also be monitored using a FACScan flow cytometer or in a single cell using fluorescence microscopy.

EXAMPLE 5 UCP3 GENE EXPRESSION: Tissue Distribution and Physiologic Regulation

Tissue Distribution - In humans, UCP3 is expressed abundantly and preferentially in skeletal muscle. In rats, UCP3 is expressed abundantly in skeletal muscle and brown fat.

Starvation - UCP3 was dramatically increased by starvation in mice and rats (~5-10 fold). In humans, it

has been shown that 5 days of food restriction causes a 2.5-fold increase in UCP3 mRNA expression. Also, it was found that human UCP3 mRNA is significantly upregulated when transgenic mice bearing a human UCP3 P1 clone are starved. Thus, it is likely that humans, like rodents, increase UCP3 gene expression with starvation.

Role of FFAs - Recently, it was shown that treatment of fed rats with Intralipid plus heparin (which produced an increase in free fatty acids (FFAs) from 0.26 to 2.04 mM) caused a 3-fold increase in UCP3 (Weigle D.S., *Diabetics*, 47:298-302 (1998)). Based upon this observation, it was suggested that the increase in FFAs with starvation was responsible for the effects of starvation on UCP3 mRNA levels. It was speculated that "this induction of UCP3 may be linked to the utilization of free fatty acids as a fuel". As discussed below however, it is unlikely that this hypothesis is true.

Starvation plus Nicotinic Acid - 1 day fasted rats were treated with saline or nicotinic acid for 6 hours and the effects on UCP3 gene expression were assessed. Starvation increases lipolysis in adipose tissue, causing a marked increase in blood levels of FFAs. The increase in FFAs is thought to promote conservation of protein in skeletal muscle (when lipid fuels are abundant, the requirement for gluconeogenesis from muscle protein is reduced). Nicotinic acid inhibits lipolysis, restores FFA levels to fed values, and stimulates protein catabolism in skeletal muscle (Lowell and Goodman, *Diabetics*, *36*:14-19 (1987)). The experiment described herein shows that nicotinic acid treatment of fasted animals returned FFA levels to fed values, but increased UCP3 mRNA to levels 2-fold higher than those observed in saline treated fasted controls. This observation shows that the starvation-induced rise in FFAs is not responsible for the effects of starvation on UCP3 mRNA levels. Also, it shows that UCP3

is not linked to the utilization of FFAs as fuel. Instead, based upon this finding it is reasonable to expect that UCP3 is linked to protein catabolism in skeletal muscle.

Streptozotocin Diabetes - Fourteen days of streptozotocin diabetes in rats produced a very large increase in UCP3 mRNA levels. This rise in UCP3 was reversed with one day of insulin treatment. Streptozotocin diabetes is associated with significant protein catabolism in skeletal muscle.

Endotoxin - Endotoxin treatment of rats and mice resulted in a very large increase in UCP3 mRNA levels in skeletal muscle, but not in other tissues. Endotoxin is a well known stimulator of protein catabolism in skeletal muscle.

Dexamethasone - High dose dexamethasone treatment markedly stimulated UCP3 mRNA levels in skeletal muscle, but not in other tissues. Dexamethasone is also a well known stimulator of protein catabolism in skeletal muscle.

Thyroid Hormone - High dose thyroid treatment in rats stimulated UCP3 mRNA levels. Thyroid hormones seemed to have little or no effect in mice. Thyroid hormone is also a well known stimulator of protein catabolism in skeletal muscle.

ob/ob and db/db mice: fa/fa rats - These genetically obese rodents were generated and shown to have markedly increased UCP3 mRNA levels in skeletal muscle. It is likely that increased UCP3 mRNA levels in ob/ob mice contributed to elevated production of gluconeogenic precursors by muscle, thereby promoting non-insulin dependent diabetes mellitus (NIDDM) in these animals.

It is interesting to note that nearly all positive regulators of UCP3 gene expression (starvation, nicotinic acid treatment during starvation, streptozotocin diabetes, endotoxin, dexamethasone and thyroid hormone) are associated with catabolism of skeletal muscle protein (see

Mitch and Goldberg, *NEJM*, 335:1897-1905 (1996)). The only exceptions to this are genetically obese rodents (however, these animals do have decreased muscle mass). From another perspective, it is also true that all catabolic states tested to date are associated with increased UCP3 expression.

Given that increased UCP3 gene expressions is linked to states of augmented skeletal muscle protein catabolism, it is likely that UCP3 plays an important role in regulating skeletal muscle protein catabolism (conversion of muscle protein to gluconeogenic precursors). Possible mechanisms by which UCP3 plays a role are the following:

a) UCP3 is a mitochondrial carrier which transports biosynthetic metabolites in and out of mitochondria during skeletal muscle protein catabolism (i.e., conversion of aspartate, glutamate, valine, isoleucine and leucine to gluconeogenic precursors alanine and glutamine).

b) UCP3 is the aspartate/glutamate carrier and is rate limiting for operation of the aspartate/malate shuttle (transfers cytosolic NADH into the mitochondria). Increased operation of this shuttle would reduce the cytosolic NADH/NAD ratio. It has been suggested that the cytosolic NADH/NAD ratio regulates muscle protein catabolism.

c) UCP3 is indeed a genuine uncoupling protein and increased UCP3 activity in catabolic states oxidizes the whole cell redox state (NADH/NAD ratio), thereby stimulating protein catabolism and amino acid metabolism.

Skeletal Muscle Metabolism During Starvation (and other catabolic states).

During starvation, muscle mobilizes actin and myosin protein and releases gluconeogenic precursors into the blood (primarily alanine and glutamine). This response is critical for survival. In the absence of gluconeogenesis from muscle protein, blood glucose levels would fall during starvation and brain dysfunction would occur.

The amino acids released from muscle protein are significantly metabolized inside the myocytes prior to their release into the bloodstream. Alanine and glutamine represent approximately 12% amino acids in muscle protein but together represent > 50% of amino acids released by muscle during starvation. Thus, much of the alanine and glutamine released must be synthesized. In contrast, aspartate, asparginine, glutamate, leucine, isoleucine and valine represent > 30% of amino acids in muscle protein but are released in only small amounts during starvation. These amino acids are interconverted to alanine and glutamine by muscle. Other amino acids such as glycine, cysteine, serine, threonine, methionine, proline, lysine, arginine, histidine, phenylalanine, tyrosine and tryptophan represents approximately 50% of muscle protein and are released either unchanged or as deaminated α-ketoacids.

Alanine is generated by the transamination of pyruvate. The pyruvate (i.e., carbon) for alanine synthesis come from glycolysis while the nitrogen originates from aspartate, asparginine, glutamate, leucine, isoleucine and valine. The released alanine is taken up by the liver and used to synthesize glucose. The glucose is then returned to the muscle and is metabolized into pyruvate, thus completing the glucose-alanine cycle. It is important to note that no new glucose is synthesized by this process, the carbons are simply recycled. Thus, the glucose-alanine cycle functions to conserve carbohydrate, but does not generate new carbohydrate. The cycle also functions to transfer NH2 from amino acids with are

metabolized (aspartate, asparginine, glutamate, leucine, isoleucine and valine) to the liver where it can be detoxified via the urea cycle.

Because certain tissues are oxidizing glucose to CO_2 (i.e., the brain), new glucose must be synthesized during starvation. This new glucose is synthesized from glutamine, which is released by muscle. The carbon backbone for glutamine comes from aspartate, asparginine, glutamate, isoleucine and valine, while the nitrogen comes from these same amino acids plus leucine. The leucine carbon backbone is completely oxidized to CO_2 by muscle. The glutamine released by muscle is taken up by the kidney and intestines, where a complex pathway is initiated culminating in the synthesis of glucose. Glutamine synthetase is the enzyme which converts glutamate to glutamine, the final step in glutamine synthesis. It is interesting to note that glutamine synthetase gene expression in muscle, like UCP3 gene expression, is induced by starvation, streptozotocin diabetes, endotoxin treatment and dexamethasone. It is also interesting to note, as was seen with UCP3, that these effects on glutamine synthetase gene expression are observed in skeletal muscle, but not in other tissues.

Significant mitochondrial metabolism must occur in order for aspartate, asparginine, glutamate, isoleucine, valine and leucine to be interconverted to alanine and glutamate. This is because important enzymes involved in the interconversion are located within the mitochondrial matrix. One example is branched chain α-ketoacid dehydrogenase (BCKADH), an enzyme which initiates the oxidation of leucine, isoleucine and valine. Of interest, BCKADH activity in muscle increases significantly during starvation and streptozotocin diabetes. Thus, metabolites

must flux in and out of mitochondria for muscle to release alanine and glutamine during catabolic states.

EQUIVALENTS

Those skilled in the art will recognize, or be able to ascertain using no more than routine experimentation, many equivalents to the specific embodiments of the invention described herein. Such equivalents are intended to be encompassed by the following claims.

SEQUENCE LISTING

(1) GENERAL INFORMATION:

(i) APPLICANT:
(A) NAME: Beth Israel Deaconess Medical Center
(B) STREET: 330 Brookline Avenue
(C) CITY: Boston
(D) STATE/PROVINCE: Massachusetts
(E) COUNTRY: USA
(F) POSTAL CODE/ZIP: 02215
(G) TELEPHONE: (617) 632-7000
(I) TELEFAX: (617) 632-7098

(ii) TITLE OF INVENTION: UPC3: AN UNCOUPLING PROTEIN HOMOLOGUE

(iii) NUMBER OF SEQUENCES: 35

(iv) CORRESPONDENCE ADDRESS:
(A) ADDRESSEE: HAMILTON, BROOK, SMITH & REYNOLDS, P.C.
(B) STREET: TWO MILITIA DRIVE
(C) CITY: LEXINGTON
(D) STATE: MASSACHUSETTS
(E) COUNTRY: USA
(F) ZIP: 02173

(v) COMPUTER READABLE FORM:
(A) MEDIUM TYPE: Floppy disk
(B) COMPUTER: IBM PC compatible
(C) OPERATING SYSTEM: PC-DOS/MS-DOS
(D) SOFTWARE: PatentIn Release #1.0, Version #1.30

(vi) CURRENT APPLICATION DATA:
(A) APPLICATION NUMBER: US 08/892,745
(B) FILING DATE: 15-JUL-1997
(C) CLASSIFICATION:

(vii) PRIOR APPLICATION DATA:
(A) APPLICATION NUMBER: US 60/046,254
(B) FILING DATE: 12-MAY-1997

(vii) PRIOR APPLICATION DATA:
(A) APPLICATION NUMBER: US 60/043,447
(B) FILING DATE: 09-APR-1997

(viii) ATTORNEY/AGENT INFORMATION:
(A) NAME: Granahan, Patricia
(B) REGISTRATION NUMBER: 32,227
(C) REFERENCE/DOCKET NUMBER: BIH97-01p2A2 PCT

(ix) TELECOMMUNICATION INFORMATION:
(A) TELEPHONE: (781) 861-6240
(B) TELEFAX: (781) 861-9540

(2) INFORMATION FOR SEQ ID NO:1:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 1220 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: other nucleic acid
(A) DESCRIPTION: /desc = "DNA"

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:1:

AGGAGGGGCC ATCCAATCCC TGCTGCCACC TCCTGGGATG GAGCCCTAGG GAGCCCCTGT 60

GCTGCCCCTG CCGTGGCAGG ACTCACAGCC CCACCGCTGC ACTGAAGCCC AGGGCTGTGG 120

AGCAGCCTCT CTCCTTGGAC CTCCTCTGGG CCCTAAAGGG ACTGGGCAGA GCCTTCCAGG 180

ACTATGGTTG GACTGAAGCC TTCAGACGTC CCTCCCACCA TGGCTGTGAA GTTCCTGGGG 240

GCAGGCACAG CAGCCTGTTT TGCTGAACTC GTTACCTTTC CACTGGACAC AGCCAAGGTC 300

CGCCTGCAGA TCCAGGGGGA GAACCAGGCG GTCCAGACGG CCCGGCTCGT GCAGTACCGT 360

GGCGTGCTGG GCACCATCCT GACCATGGTG CGGACTGAGG GTCCCTGCAG CCCCTACAAT 420

GGGCTGGTGG CCGGCCTGCA GCGCCAGATG AGCTTCGCCT CCATCCGCAT CGGCCTCTAT 480

GACTCCGTCA AGCAGGTGTA CACCCCCAAA GGCGCGGACA ACTCCAGCCT CACTACCCGG 540

ATTTTGGCCG GCTGCACCAC AGGAGCCATG GCGGTGACCT GTGCCCAGCC CACAGATGTG 600

GTGAAGGTCC GATTTCAGGC CAGCATACAC CTCGGGCCAT CCAGGAGCGA CAGAAAATAC 660

AGCGGGACTA TGGACGCCTA CAGAACCATC GCCAGGGAGG AAGGAGTCAG GGGCCTGTGG 720

AAAGGAACTT TGCCCAACAT CATGAGGAAT GCTATCGTCA ACTGTGCTGA GGTGGTGACC 780

TACGACATCC TCAAGGAGAA GCTGCTGGAC TACCACCTGC TCACTGACAA CTTCCCCTGC 840

CACTTTGTCT CTGCCTTTGG AGCCGGCTTC TGTGCCACAG TGGTGGCCTC CCCGGTGGAC 900

GTGGTGAAGA CCCGGTATAT GAACTCACCT CCAGGCCAGT ACTTCAGCCC CCTCGACTGT 960

ATGATAAAGA TGGTGGCCCA GGAGGGCCCC ACAGCCTTCT ACAAGGGATT TACACCCTCC 1020

TTTTTGCGTT TGGGATCCTG GAACGTGGTG ATGTTCGTAA CCTATGAGCA GCTGAAACGG 1080

GCCCTGATGA AAGTCCAGAT GTTACGGGAA TCACCGTTTT GAACAAGACA AGAAGGCCAC 1140

TGGTAGCTAA CGTGTCCGAA ACCAGTTAAG AATGGAAGAA AACGGTGCAT CCACGCACAC 1200

ATGGACACAG ACCCACACAT 1220

(2) INFORMATION FOR SEQ ID NO:2:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 1034 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: other nucleic acid
(A) DESCRIPTION: /desc = "DNA"

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:2:

GAGGGGCCAT CCAATCCCTG CTGCCACCTC CTGGGATGGA GCCCTAGGGA GCCCCTGTGC 60

TGCCCCTGCC GTGGCAGGAC TCACAGCCCC ACCGCTGCCT GAAGCCCAGG GCTGTGGAGC 120

AGCCTCTCTC CTTGGACCTC CTCTCGGCCC TAAAGGGACT GGGCAGAGCC TTCCAGGACT 180

ATGGTTGGAC TGAAGCCTTC AGACGTGCCT CCCACCATGG CTGTGAAGTT CCTGGGGGCA 240

GGCACAGCAG CCTGTTTTGC TGAACTCGTT ACCTTTCCAC TGGACACAGC CAAGGTCCGC 300

CTGCAGATCC AGGGGGAGAA CCAGGCGGTC CAGACGGCCC GGCTCGTGCA GTACCGTGGC 360

GTGCTGGGCA CCATCCTGAC CATGGTGCGG ACTGAGGGTC CCTGCAGCCC CTACAATGGG 420

CTGGTGGCCG GCCTGCAGCG CCAGATGAGC TTCGCCTCCA TCCGCATCGG CCTCTATGAC 480

TCCGTCAAGC AGGTGTACAC CCCCAAAGGC GCGGACAACT TCCAGCCTCA CTACCCGGAT 540

TTTGGCCGGC TGCACCACAG GAGCCATGGC GGTGACCTGT GCCCAGCCCA CAGATGTGGT 600

GAAGGTCCGA TTTCAGGCCA GCATACACCT CGGGCCATCC AGGACCGACA GAAAATACAG 660

CGGGACTATG GACGCCTACA GAACCATCGC CAGGGAGGAA GGAGTCAGGG GCCTGTGGAA 720

AGGAACTTTG CCCAACATCA TGAGGAATGC TATCGTCAAC TGTGCTGAGG TGGTGACCTA 780

CGACATCCTC AAGGAGAAGC TGCTGGACTA CCACCTGCTC ACTGACAACT TCCCCTGCCA 840

CTTTGTCTCT GCCTTTGGAG CCGGCTTCTG TGCCACAGTG GTGGCCTCCC CGGTGGACGT 900

GGTGAAGACC CGGTATATGA ACTCACCTCC AGGCCAGTAC TTCAGCCCCC TCGACTGTAT 960

GATAAAGATG GTGGCCCAGG AGGGCCCCAC AGCCTTCTAC AAGGGGTGAG CCTCCTCCTG 1020

CCTCCAGCAC TCCC 1034

(2) INFORMATION FOR SEQ ID NO:3:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 312 amino acids
(B) TYPE: amino acid
(C) STRANDEDNESS: unknown
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:3:

```
Met Val Gly Leu Lys Pro Ser Asp Val Pro Pro Thr Met Ala Val Lys
1               5                  10                  15

Phe Leu Gly Ala Gly Thr Ala Ala Cys Phe Ala Glu Leu Val Thr Phe
            20                  25                  30

Pro Leu Asp Thr Ala Lys Val Arg Leu Gln Ile Gln Gly Glu Asn Gln
        35                  40                  45

Ala Val Gln Thr Ala Arg Leu Val Gln Tyr Arg Gly Val Leu Gly Thr
    50                  55                  60

Ile Leu Thr Met Val Arg Thr Glu Gly Pro Cys Ser Pro Tyr Asn Gly
65                  70                  75                  80

Leu Val Ala Gly Leu Gln Arg Gln Met Ser Phe Ala Ser Ile Arg Ile
                85                  90                  95

Gly Leu Tyr Asp Ser Val Lys Gln Val Tyr Thr Pro Lys Gly Ala Asp
            100                 105                 110

Asn Ser Ser Leu Thr Thr Arg Ile Leu Ala Gly Cys Thr Thr Gly Ala
        115                 120                 125

Met Ala Val Thr Cys Ala Gln Pro Thr Asp Val Val Lys Val Arg Phe
    130                 135                 140

Gln Ala Ser Ile His Leu Gly Pro Ser Arg Ser Asp Arg Lys Tyr Ser
145                 150                 155                 160

Gly Thr Met Asp Ala Tyr Arg Thr Ile Ala Arg Glu Glu Gly Val Arg
                165                 170                 175

Gly Leu Trp Lys Gly Thr Leu Pro Asn Ile Met Arg Asn Ala Ile Val
            180                 185                 190

Asn Cys Ala Glu Val Val Thr Tyr Asp Ile Leu Lys Glu Lys Leu Leu
        195                 200                 205

Asp Tyr His Leu Leu Thr Asp Asn Phe Pro Cys His Phe Val Ser Ala
    210                 215                 220

Phe Gly Ala Gly Phe Cys Ala Thr Val Val Ala Ser Pro Val Asp Val
225                 230                 235                 240

Val Lys Thr Arg Tyr Met Asn Ser Pro Pro Gly Gln Tyr Phe Ser Pro
                245                 250                 255

Leu Asp Cys Met Ile Lys Met Val Ala Gln Glu Gly Pro Thr Ala Phe
            260                 265                 270

Tyr Lys Gly Phe Thr Pro Ser Phe Leu Arg Leu Gly Ser Trp Asn Val
        275                 280                 285

Val Met Phe Val Thr Tyr Glu Gln Leu Lys Arg Ala Leu Met Lys Val
    290                 295                 300

Gln Met Leu Arg Glu Ser Pro Phe
305                 310
```

(2) INFORMATION FOR SEQ ID NO:4:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 275 amino acids
(B) TYPE: amino acid
(C) STRANDEDNESS: unknown
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:4:

Met Val Gly Leu Lys Pro Ser Asp Val Pro Pro Thr Met Ala Val Lys
1 5 10 15

Phe Leu Gly Ala Gly Thr Ala Ala Cys Phe Ala Glu Leu Val Thr Phe
20 25 30

Pro Leu Asp Thr Ala Lys Val Arg Leu Gln Ile Gln Gly Glu Asn Gln
35 40 45

Ala Val Gln Thr Ala Arg Leu Val Gln Tyr Arg Gly Val Leu Gly Thr
50 55 60

Ile Leu Thr Met Val Arg Thr Glu Gly Pro Cys Ser Pro Tyr Asn Gly
65 70 75 80

Leu Val Ala Gly Leu Gln Arg Gln Met Ser Phe Ala Ser Ile Arg Ile
85 90 95

Gly Leu Tyr Asp Ser Val Lys Gln Val Tyr Thr Pro Lys Gly Ala Asp
100 105 110

Asn Ser Ser Leu Thr Thr Arg Ile Leu Ala Gly Cys Thr Thr Gly Ala
115 120 125

Met Ala Val Thr Cys Ala Gln Pro Thr Asp Val Val Lys Val Arg Phe
130 135 140

Gln Ala Ser Ile His Leu Gly Pro Ser Arg Ser Asp Arg Lys Tyr Ser
145 150 155 160

Gly Thr Met Asp Ala Tyr Arg Thr Ile Ala Arg Glu Glu Gly Val Arg
165 170 175

Gly Leu Trp Lys Gly Thr Leu Pro Asn Ile Met Arg Asn Ala Ile Val
180 185 190

Asn Cys Ala Glu Val Val Thr Tyr Asp Ile Leu Lys Glu Lys Leu Leu
195 200 205

Asp Tyr His Leu Leu Thr Asp Asn Phe Pro Cys His Phe Val Ser Ala
210 215 220

Phe Gly Ala Gly Phe Cys Ala Thr Val Val Ala Ser Pro Val Asp Val
225 230 235 240

Val Lys Thr Arg Tyr Met Asn Ser Pro Pro Gly Gln Tyr Phe Ser Pro
245 250 255

Leu Asp Cys Met Ile Lys Met Val Ala Gln Glu Gly Pro Thr Ala Phe
260 265 270

Tyr Lys Gly
275

(2) INFORMATION FOR SEQ ID NO:5:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 307 amino acids
(B) TYPE: amino acid
(C) STRANDEDNESS: unknown
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:5:

Met Gly Gly Leu Thr Ala Ser Asp Val His Pro Thr Leu Gly Val Gln
1 5 10 15

Leu Phe Ser Ala Gly Ile Ala Ala Cys Leu Ala Asp Val Ile Thr Phe
20 25 30

Pro Leu Asp Thr Ala Lys Val Arg Leu Gln Val Gln Gly Glu Cys Pro
35 40 45

Thr Ser Ser Val Ile Arg Tyr Lys Gly Val Leu Gly Thr Ile Thr Ala
50 55 60

Val Val Lys Thr Glu Gly Arg Met Lys Leu Tyr Ser Gly Leu Pro Ala
65 70 75 80

Gly Leu Gln Arg Gln Ile Ser Ser Ala Ser Leu Arg Ile Gly Leu Tyr
85 90 95

Asp Thr Val Gln Glu Phe Leu Thr Ala Gly Lys Glu Thr Ala Pro Ser
100 105 110

Leu Gly Ser Lys Ile Leu Ala Gly Leu Thr Thr Gly Gly Val Ala Val
115 120 125

Phe Ile Gly Gln Pro Thr Glu Val Val Lys Val Arg Leu Gln Ala Gln
130 135 140

Ser His Leu His Gly Ile Lys Pro Arg Tyr Thr Gly Thr Tyr Asn Ala
145 150 155 160

Tyr Arg Ile Ile Ala Thr Thr Glu Gly Leu Thr Gly Leu Trp Lys Gly
165 170 175

Thr Thr Pro Asn Leu Met Arg Ser Val Ile Ile Asn Cys Thr Glu Leu
180 185 190

Val Thr Tyr Asp Leu Met Lys Glu Ala Phe Val Lys Asn Asn Ile Leu
195 200 205

Ala Asp Asp Val Pro Cys His Leu Val Ser Ala Leu Ile Ala Gly Phe
210 215 220

Cys Ala Thr Ala Met Ser Ser Pro Val Asp Val Val Lys Thr Arg Phe
225 230 235 240

Ile Asn Ser Pro Pro Gly Gln Tyr Lys Ser Val Pro Asn Cys Ala Met
245 250 255

Lys Val Phe Thr Asn Glu Gly Pro Thr Ala Phe Phe Lys Gly Leu Val
260 265 270

Pro Ser Phe Leu Arg Leu Gly Ser Trp Asn Val Ile Met Phe Val Cys
275 280 285

Phe Glu Gln Leu Lys Arg Glu Leu Ser Lys Ser Arg Gln Thr Met Asp
290 295 300

Cys Ala Thr
305

(2) INFORMATION FOR SEQ ID NO:6:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 308 amino acids
(B) TYPE: amino acid
(C) STRANDEDNESS: unknown
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:6:

Met Val Gly Phe Lys Ala Thr Asp Val Pro Pro Thr Ala Thr Val Lys
1 5 10 15

Leu Phe Gly Ala Gly Thr Ala Ala Cys Ile Ala Asp Leu Ile Thr Phe
20 25 30

Pro Leu Asp Thr Ala Lys Val Arg Leu Gln Ile Gln Gly Glu Ser Gln
35 40 45

Gly Pro Val Arg Ala Thr Val Ser Ala Gln Tyr Arg Gly Val Met Gly
50 55 60

Thr Ile Leu Thr Met Val Arg Thr Glu Gly Pro Arg Ser Leu Tyr Asn
65 70 75 80

Cys Leu Val Ala Gly Leu Gln Arg Gln Met Ser Phe Ala Ser Val Arg
85 90 95

Ile Gly Leu Tyr Asp Ser Val Lys Gln Phe Tyr Thr Lys Gly Ser Glu
100 105 110

His Ala Ser Ile Gly Ser Arg Leu Leu Ala Gly Ser Thr Thr Gly Ala
115 120 125

Leu Ala Val Ala Val Ala Gln Pro Thr Asp Val Val Lys Val Arg Phe
130 135 140

Gln Ala Gln Arg Ala Gly Gly Gly Arg Arg Tyr Gln Ser Thr Val Asn
145 150 155 160

Ala Tyr Lys Thr Ile Ala Arg Glu Glu Gly Phe Arg Gly Leu Trp Lys
165 170 175

Gly Thr Ser Pro Asn Val Ala Arg Asn Ala Ile Val Asn Cys Ala Glu
180 185 190

Leu Val Thr Tyr Asp Leu Ile Lys Asp Ala Leu Leu Lys Ala Asn Leu
195 200 205

Met Thr Asp Asp Leu Pro Cys His Phe Thr Ser Ala Phe Gly Ala Gly
210 215 220

Phe Cys Thr Thr Val Ile Ala Ser Pro Val Asp Val Val Lys Thr Arg
225 230 235 240

Tyr Met Asn Ser Ala Leu Gly Gln Tyr Ser Ser Ala Gly His Cys Ala
245 250 255

Leu Thr Met Leu Gln Lys Glu Gly Pro Arg Ala Phe Tyr Lys Gly Phe
260 265 270

Met Pro Ser Phe Leu Arg Leu Gly Ser Trp Asn Val Val Met Phe Val
275 280 285

Thr Tyr Glu Gln Leu Lys Arg Ala Leu Met Ala Ala Cys Thr Ser Arg
290 295 300

Glu Ala Pro Phe
305

(2) INFORMATION FOR SEQ ID NO:7:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 1204 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: other nucleic acid
(A) DESCRIPTION: /desc = "DNA"

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:7:

GAGACAACAG TGAATGGTGA GGCCCGGCCG TCAGATCCTG CTGCTACCTA ATGGAGTGGA 60

GCCTTAGGGT GGCCCTGCAC TACCCAACCT TGGCTAGACG CACAGCTTCC TCCCTGAACT 120

GAAGCAAAAG ATTGCCAGGC AAGCTCTCTC CTCGGACCTC CATAGGCAGC AAAGGAACCA 180

GGCCCATTCC CCGGGACCAT GGTTGGACTT CAGCCCTCCG AAGTGCCTCC CACAACGGTT 240

GTGAAGTTCC TGGGGGCCGG CACTGCGGCC TGTTTTGCGG ACCTCCTCAC TTTTCCCCTG 300

GACACCGCCA AGGTCCGTCT GCAGATCCAA GGGGAGAACC CAGGGGCTCA GAGCGTGCAG 360

TACCGCGGTG TGCTGGGTAC CATCCTGACT ATGGTGCGCA CAGAGGGTCC CCGCAGCCCC 420

TACAGCGGAC TGGTCGCTGG CCTGCACCGC CAGATGAGTT TTGCCTCCAT TCGAATTGGC 480

CTCTACGACT CTGTCAAGCA GTTCTACACC CCCAAGGGAG CGGACCACTC CAGCGTCGCC 540

ATCAGGATTC TGGCAGGCTG CACGACAGGA GCCATGGCAG TGACCTGCGC CCAGCCCACG 600

GATGTGGTCA AGGTCCGATT TCAAGCCATG ATACGCCTGG GAACTGGAGG AGAGAGGAAA 660

TACAGAGGGA CTATGGATGC CTACAGAACC ATCGCCAGGG AGGAAGGAGT CAGGGGCCTG 720

TGGAAAGGGA CTTGGCCCAA CATCACAAGA AATGCCATTG TCAACTGTGC TGAGATGGTG 780

ACCTACGACA TCATCAAGGA GAAGTTGCTG GAGTCTCACC TGTTTACTGA CAACTTCCCC 840

TGTCACTTTG TCTCTGCCTT TGGAGCTGGC TTCTGTGCCA CAGTGGTGGC CTCCCCGGTC 900

GATGTGGTAA AGACCCGATA CATGAACGCT CCCCTAGGCA GGTACCGCAG CCCTCTGCAC 960

TGTATGCTGA AGATGGTGGC TCACGAGGGA CCCACGGCCT TCTACAAAGG ATTTGTGCCC 1020

TCCTTTCTGC GTCTGGGAGC TTGGAACGTG ATGATGTTTG TAACATATCA GCAACTGAAG 1080

AGGGCCTTAA TGAAAGTCCA GGTACTGCGG GAATCTCCGT TTTGAACAAG GCAAGCAGGC 1140

TGCCTGGAAC AGAACAAAGC GTCTCTGCCT GGGACACAGG CCCACACGTC AGAACCGTGC 1200

ACGC 1204

(2) INFORMATION FOR SEQ ID NO:8:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 308 amino acids
(B) TYPE: amino acid
(C) STRANDEDNESS: unknown
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:8:

Met Val Gly Leu Gln Pro Ser Glu Val Pro Pro Thr Thr Val Val Lys
1 5 10 15

Phe Leu Gly Ala Gly Thr Ala Ala Cys Phe Ala Asp Leu Leu Thr Phe
20 25 30

Pro Leu Asp Thr Ala Lys Val Arg Leu Gln Ile Gln Gly Glu Asn Pro
35 40 45

Gly Ala Gln Ser Val Gln Tyr Arg Gly Val Leu Gly Thr Ile Leu Thr
50 55 60

Met Val Arg Thr Glu Gly Pro Arg Ser Pro Tyr Ser Gly Leu Val Ala
65 70 75 80

Gly Leu His Arg Gln Met Ser Phe Ala Ser Ile Arg Ile Gly Leu Tyr
85 90 95

Asp Ser Val Lys Gln Phe Tyr Thr Pro Lys Gly Ala Asp His Ser Ser
100 105 110

Val Ala Ile Arg Ile Leu Ala Gly Cys Thr Thr Gly Ala Met Ala Val
115 120 125

```
Thr Cys Ala Gln Pro Thr Asp Val Val Lys Val Arg Phe Gln Ala Met
    130                 135                 140

Ile Arg Leu Gly Thr Gly Gly Glu Arg Lys Tyr Arg Gly Thr Met Asp
145                 150                 155                 160

Ala Tyr Arg Thr Ile Ala Arg Glu Glu Gly Val Arg Gly Leu Trp Lys
                165                 170                 175

Gly Thr Trp Pro Asn Ile Thr Arg Asn Ala Ile Val Asn Cys Ala Glu
            180                 185                 190

Met Val Thr Tyr Asp Ile Ile Lys Glu Lys Leu Leu Glu Ser His Leu
        195                 200                 205

Phe Thr Asp Asn Phe Pro Cys His Phe Val Ser Ala Phe Gly Ala Gly
    210                 215                 220

Phe Cys Ala Thr Val Val Ala Ser Pro Val Asp Val Val Lys Thr Arg
225                 230                 235                 240

Tyr Met Asn Ala Pro Leu Gly Arg Tyr Arg Ser Pro Leu His Cys Met
                245                 250                 255

Leu Lys Met Val Ala Gln Glu Gly Pro Thr Ala Phe Tyr Lys Gly Phe
            260                 265                 270

Val Pro Ser Phe Leu Arg Leu Gly Ala Trp Asn Val Met Met Phe Val
        275                 280                 285

Thr Tyr Glu Gln Leu Lys Arg Ala Leu Met Lys Val Gln Val Leu Arg
    290                 295                 300

Glu Ser Pro Phe
305
```

(2) INFORMATION FOR SEQ ID NO:9:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 307 amino acids
(B) TYPE: amino acid
(C) STRANDEDNESS: unknown
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:9:

```
Met Val Asn Pro Thr Thr Ser Glu Val Gln Pro Thr Met Gly Val Lys
1               5                   10                  15

Ile Phe Ser Ala Gly Val Ser Ala Cys Leu Ala Asp Ile Ile Thr Phe
            20                  25                  30

Pro Leu Asp Thr Ala Lys Val Arg Leu Gln Ile Gln Gly Glu Gly Gln
        35                  40                  45
```

Ala Ser Ser Thr Ile Arg Tyr Lys Gly Val Leu Gly Thr Ile Thr Thr
50 55 60

Leu Ala Lys Thr Glu Gly Leu Pro Lys Leu Tyr Ser Gly Leu Pro Ala
65 70 75 80

Gly Ile Gln Arg Gln Ile Ser Phe Ala Ser Leu Arg Ile Gly Leu Tyr
85 90 95

Asp Ser Val Gln Glu Tyr Phe Ser Ser Gly Arg Glu Thr Pro Ala Ser
100 105 110

Leu Gly Asn Lys Ile Ser Ala Gly Leu Met Thr Gly Gly Val Ala Val
115 120 125

Phe Ile Gly Gln Pro Thr Glu Val Val Lys Val Arg Met Gln Ala Gln
130 135 140

Ser His Leu His Gly Ile Lys Pro Arg Tyr Thr Gly Thr Tyr Asn Ala
145 150 155 160

Tyr Arg Val Ile Ala Thr Thr Glu Ser Leu Ser Thr Leu Trp Lys Gly
165 170 175

Thr Thr Pro Asn Leu Met Arg Asn Val Ile Ile Asn Cys Thr Glu Leu
180 185 190

Val Thr Tyr Asp Leu Met Lys Gly Ala Leu Val Asn Asn Lys Ile Leu
195 200 205

Ala Asp Asp Val Pro Cys His Leu Leu Ser Ala Leu Val Ala Gly Phe
210 215 220

Cys Thr Thr Leu Leu Ala Ser Pro Val Asp Val Val Lys Thr Arg Phe
225 230 235 240

Ile Asn Ser Leu Pro Gly Gln Tyr Pro Ser Val Pro Ser Cys Ala Met
245 250 255

Ser Met Tyr Thr Lys Glu Gly Pro Thr Ala Phe Phe Lys Gly Phe Val
260 265 270

Ala Ser Phe Leu Arg Leu Gly Ser Trp Asn Val Ile Met Phe Val Cys
275 280 285

Phe Glu Gln Leu Lys Lys Glu Leu Met Lys Ser Arg Gln Thr Val Asp
290 295 300

Cys Thr Thr
305

(2) INFORMATION FOR SEQ ID NO:10:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 309 amino acids
(B) TYPE: amino acid
(C) STRANDEDNESS: unknown
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:10:

Met Val Gly Phe Lys Ala Thr Asp Val Pro Pro Thr Ala Thr Val Lys
1 5 10 15

Phe Leu Gly Ala Gly Thr Ala Ala Cys Ile Ala Asp Leu Ile Thr Phe
20 25 30

Pro Leu Asp Thr Ala Lys Val Arg Leu Gln Ile Gln Gly Glu Ser Gln
35 40 45

Gly Leu Val Arg Thr Ala Ala Ser Ala Gln Tyr Arg Gly Val Leu Gly
50 55 60

Thr Ile Leu Thr Met Val Arg Thr Glu Gly Pro Arg Ser Leu Tyr Asn
65 70 75 80

Gly Leu Val Ala Gly Leu Gln Arg Gln Met Ser Phe Ala Ser Val Arg
85 90 95

Ile Gly Leu Tyr Asp Ser Val Lys Gln Phe Tyr Thr Lys Gly Ser Glu
100 105 110

His Ala Gly Ile Gly Ser Arg Leu Leu Ala Gly Ser Thr Thr Gly Ala
115 120 125

Leu Ala Val Ala Val Ala Gln Pro Thr Asp Val Val Lys Val Arg Phe
130 135 140

Gln Ala Gln Ala Arg Ala Gly Gly Gly Arg Arg Tyr Gln Ser Thr Val
145 150 155 160

Glu Ala Tyr Lys Thr Ile Ala Arg Glu Glu Gly Ile Arg Gly Leu Trp
165 170 175

Lys Gly Thr Ser Pro Asn Val Ala Arg Asn Ala Ile Val Asn Cys Ala
180 185 190

Glu Leu Val Thr Tyr Asp Leu Ile Lys Asp Thr Leu Leu Lys Ala Asn
195 200 205

Leu Met Thr Asp Asp Leu Pro Cys His Phe Thr Ser Ala Phe Gly Ala
210 215 220

Gly Phe Cys Thr Thr Val Ile Ala Ser Pro Val Asp Val Val Lys Thr
225 230 235 240

Arg Tyr Met Asn Ser Ala Leu Gly Gln Tyr His Ser Ala Gly His Cys
245 250 255

Ala Leu Thr Met Ile Arg Lys Glu Gly Pro Arg Ala Phe Tyr Lys Gly
260 265 270

Phe Met Pro Ser Phe Leu Arg Leu Gly Ser Trp Asn Val Val Met Phe
275 280 285

Val Thr Tyr Glu Gln Leu Lys Arg Ala Leu Met Ala Ala Tyr Gln Ser
290 295 300

Arg Glu Ala Pro Phe
305

(2) INFORMATION FOR SEQ ID NO:11:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 20 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: other nucleic acid
(A) DESCRIPTION: /desc = "DNA"

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:11:

GGACTCACAG GTAAGACCCC 20

(2) INFORMATION FOR SEQ ID NO:12:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 20 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: other nucleic acid
(A) DESCRIPTION: /desc = "DNA"

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:12:

TCTCCTGCAG CCCCACCGCT 20

(2) INFORMATION FOR SEQ ID NO:13:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 20 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: other nucleic acid
(A) DESCRIPTION: /desc = "DNA"

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:13:

CCGCCTGCAG GTAGGTGCCC 20

(2) INFORMATION FOR SEQ ID NO:14:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 20 amino acids
(B) TYPE: amino acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: other nucleic acid

(A) DESCRIPTION: /desc = "DNA"

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:14:

Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Ala Thr Cys Cys Ala Gly
1 5 10 15

Gly Gly Gly Gly
20

(2) INFORMATION FOR SEQ ID NO:15:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 20 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: other nucleic acid
(A) DESCRIPTION: /desc = "DNA"

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:15:

GGCGCGGACA GTGAGTGACC 20

(2) INFORMATION FOR SEQ ID NO:16:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 20 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: other nucleic acid
(A) DESCRIPTION: /desc = "DNA"

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:16:

CCCCTCCCAG ACTCCAGCCT 20

(2) INFORMATION FOR SEQ ID NO:17:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 20 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: other nucleic acid
(A) DESCRIPTION: /desc = "DNA"

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:17:

CTGTGGAAAG GTAGGTCTGG 20

(2) INFORMATION FOR SEQ ID NO:18:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 20 amino acids
(B) TYPE: amino acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: other nucleic acid
(A) DESCRIPTION: /desc = "DNA"

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:18:

Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Gly Ala Ala Cys Thr Thr
1 5 10 15

Thr Gly Cys Cys
20

(2) INFORMATION FOR SEQ ID NO:19:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 20 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: other nucleic acid
(A) DESCRIPTION: /desc = "DNA"

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:19:

CTGCTCACTG GTGAGGCCCT 20

(2) INFORMATION FOR SEQ ID NO:20:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 20 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: other nucleic acid
(A) DESCRIPTION: /desc = "DNA"

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:20:

TCCTCTGCAG ACAACTTCCC 20

(2) INFORMATION FOR SEQ ID NO:21:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 20 base pairs
(B) TYPE: nucleic acid

(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: other nucleic acid
(A) DESCRIPTION: /desc = "DNA"

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:21:

TCTACAAGGG GTGAGCCTCC 20

(2) INFORMATION FOR SEQ ID NO:22:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 20 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: other nucleic acid
(A) DESCRIPTION: /desc = "DNA"

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:22:

TTCTTATCAG ATTTACACCC 20

(2) INFORMATION FOR SEQ ID NO:23:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 20 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: other nucleic acid
(A) DESCRIPTION: /desc = "DNA"

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:23:

GGACTACCAC CTGCTCACTG 20

(2) INFORMATION FOR SEQ ID NO:24:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 19 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: other nucleic acid
(A) DESCRIPTION: /desc = "DNA"

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:24:

CCCGTAACAT ATGGACTTT 19

(2) INFORMATION FOR SEQ ID NO:25:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 30 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: other nucleic acid
(A) DESCRIPTION: /desc = "DNA"

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:25:

TTCACCACGT CCACCCGGGG GGATGCCACC 30

(2) INFORMATION FOR SEQ ID NO:26:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 6 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: other nucleic acid
(A) DESCRIPTION: /desc = "DNA"

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:26:

AATAAA 6

(2) INFORMATION FOR SEQ ID NO:27:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 403 amino acids
(B) TYPE: amino acid
(C) STRANDEDNESS: unknown
(D) TOPOLOGY: unknown

(ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:27:

Arg Arg Gly His Pro Ile Pro Ala Ala Thr Ser Trp Asp Gly Ala Leu
1 5 10 15

Gly Ser Pro Cys Ala Ala Pro Ala Val Ala Gly Ile Thr Ala Pro Pro
20 25 30

Leu His Ser Pro Gly Leu Trp Ser Ser Leu Ser Pro Trp Thr Ser Ser
35 40 45

```
Arg Pro Arg Asp Trp Ala Glu Pro Ser Arg Thr Met Val Gly Leu Lys
    50                  55                  60

Pro Ser Asp Val Pro Pro Thr Met Ala Val Lys Phe Leu Gly Ala Gly
65                  70                  75                  80

Ile Ala Ala Cys Phe Ala Glu Leu Val Thr Phe Pro Leu Asp Thr Ala
                85                  90                  95

Lys Val Arg Leu Gln Ile Gln Gly Glu Asn Gln Ala Val Gln Thr Ala
            100                 105                 110

Arg Leu Val Gln Tyr Arg Gly Val Leu Gly Thr Ile Leu Thr Met Val
        115                 120                 125

Arg Thr Glu Gly Pro Cys Ser Pro Tyr Asn Gly Leu Val Ala Gly Leu
    130                 135                 140

Gln Arg Gln Met Ser Phe Ala Ser Ile Arg Ile Gly Leu Tyr Asp Ser
145                 150                 155                 160

Val Lys Gln Val Tyr Thr Pro Lys Gly Ala Asp Asn Ser Ser Leu Thr
                165                 170                 175

Thr Arg Ile Leu Ala Gly Cys Thr Thr Gly Ala Met Ala Val Thr Cys
            180                 185                 190

Ala Gln Pro Thr Asp Val Val Lys Val Arg Phe Gln Ala Ser Ile His
        195                 200                 205

Leu Gly Pro Ser Arg Ser Asp Arg Lys Tyr Ser Gly Thr Met Asp Ala
    210                 215                 220

Tyr Arg Thr Ile Ala Arg Glu Glu Gly Val Arg Gly Leu Trp Lys Gly
225                 230                 235                 240

Thr Leu Pro Asn Ile Met Arg Asn Ala Ile Val Asn Cys Ala Glu Val
                245                 250                 255

Val Thr Tyr Asp Ile Leu Lys Glu Lys Leu Leu Asp Tyr His Leu Leu
            260                 265                 270

Thr Asp Asn Phe Pro Cys His Phe Val Ser Ala Phe Gly Ala Gly Phe
        275                 280                 285

Cys Ala Thr Val Val Ala Ser Pro Val Asp Val Val Lys Thr Arg Tyr
    290                 295                 300

Met Asn Ser Pro Pro Gly Gln Tyr Phe Ser Pro Leu Asp Cys Met Ile
305                 310                 315                 320

Lys Met Val Ala Gln Glu Gly Pro Thr Ala Phe Tyr Lys Gly Phe Thr
                325                 330                 335

Pro Ser Phe Leu Arg Leu Gly Ser Trp Asn Val Val Met Phe Val Thr
            340                 345                 350

Tyr Glu Gln Leu Lys Arg Ala Leu Met Lys Val Gln Met Leu Arg Glu
        355                 360                 365

Ser Pro Phe Tyr Arg Gln Glu Gly His Trp Leu Thr Cys Pro Lys Pro
    370                 375                 380
```

Val Lys Asn Gly Arg Lys Arg Cys Ile His Ala His Met Asp Thr Asp
385 390 395 400

Pro His Ile

(2) INFORMATION FOR SEQ ID NO:28:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 387 amino acids
(B) TYPE: amino acid
(C) STRANDEDNESS: unknown
(D) TOPOLOGY: unknown

(ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:28:

Gly Gly Ala Ile Gln Ser Leu Leu Pro Pro Pro Gly Met Glu Pro Gly
1 5 10 15

Ala Pro Val Leu Pro Leu Pro Trp Gln Asp Ser Gln Pro His Arg Cys
20 25 30

Thr Glu Ala Gln Gly Cys Gly Ala Ala Ser Leu Leu Gly Pro Pro Leu
35 40 45

Gly Pro Lys Gly Thr Gly Gln Ser Leu Pro Gly Leu Trp Leu Asp Ser
50 55 60

Leu Gln Thr Cys Leu Pro Pro Trp Leu Ser Ser Trp Gly Gln Ala Gln
65 70 75 80

Gln Pro Val Leu Leu Asn Ser Leu Pro Phe His Trp Thr Gln Pro Arg
85 90 95

Ser Ala Cys Arg Ser Arg Gly Arg Thr Arg Arg Ser Arg Arg Pro Gly
100 105 110

Ser Cys Ser Thr Val Ala Cys Trp Ala Pro Ser Pro Trp Cys Gly Leu
115 120 125

Arg Val Pro Ala Ala Pro Thr Met Gly Trp Trp Pro Ala Cys Ser Ala
130 135 140

Arg Ala Ser Pro Pro Ser Ala Ser Ala Ser Met Thr Pro Ser Ser Arg
145 150 155 160

Cys Thr Pro Pro Lys Ala Arg Thr Thr Pro Ala Ser Leu Pro Gly Phe
165 170 175

Trp Pro Ala Ala Pro Gln Glu Pro Trp Arg Pro Val Pro Ser Pro Gln
180 185 190

Met Trp Arg Ser Asp Phe Arg Pro Ala Tyr Thr Ser Gly His Pro Gly
195 200 205

```
Ala Thr Glu Asn Thr Ala Gly Leu Trp Thr Pro Thr Glu Pro Ser Pro
    210                 215                 220

Gly Arg Lys Glu Ser Gly Ala Cys Gly Lys Glu Leu Cys Pro Thr Ser
225                 230                 235                 240

Gly Met Leu Ser Ser Thr Val Leu Arg Trp Pro Thr Thr Ser Ser Arg
                245                 250                 255

Arg Ser Cys Trp Thr Thr Thr Cys Ser Leu Thr Thr Ser Pro Ala Thr
            260                 265                 270

Leu Ser Leu Pro Leu Glu Pro Ala Ser Val Pro Gln Trp Trp Pro Pro
        275                 280                 285

Arg Trp Thr Trp Arg Pro Gly Ile Thr His Leu Gln Ala Ser Thr Ser
    290                 295                 300

Ala Pro Ser Thr Val Arg Trp Trp Pro Arg Arg Ala Pro Gln Pro Ser
305                 310                 315                 320

Thr Arg Asp Leu His Pro Pro Phe Cys Val Trp Asp Pro Gly Thr Trp
                325                 330                 335

Cys Ser Pro Met Ser Ser Asn Gly Pro Lys Ser Arg Cys Tyr Gly Asn
            340                 345                 350

His Arg Phe Glu Gln Asp Lys Lys Ala Thr Gly Ser Arg Val Arg Asn
        355                 360                 365

Gln Leu Arg Met Glu Glu Asn Gly Ala Ser Thr His Thr Trp Thr Gln
    370                 375                 380

Thr His Thr
385
```

(2) INFORMATION FOR SEQ ID NO:29:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 399 amino acids
(B) TYPE: amino acid
(C) STRANDEDNESS: unknown
(D) TOPOLOGY: unknown

(ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:29:

```
Glu Gly Pro Ser Asn Pro Cys Cys His Leu Leu Gly Trp Ser Pro Arg
1               5                   10                  15

Glu Pro Leu Cys Cys Pro Cys Arg Gly Arg Thr His Ser Pro Thr Ala
            20                  25                  30

Ala Leu Lys Pro Arg Ala Val Glu Gln Pro Leu Ser Leu Asp Leu Leu
        35                  40                  45

Ser Ala Leu Lys Gly Leu Gly Arg Ala Phe Gln Gln Tyr Gly Trp Thr
    50                  55                  60
```

Glu Ala Phe Arg Arg Ala Ser His His Gly Cys Glu Val Pro Gly Gly
65 70 75 80

Arg His Ser Ser Leu Phe Cys Thr Arg Tyr Leu Ser Thr Gly His Ser
85 90 95

Gln Gly Pro Pro Ala Asp Pro Gly Gly Glu Pro Gly Gly Pro Gln Gly
100 105 110

Pro Ala Arg Ala Val Pro Trp Arg Ala Gly His His Pro Asp His Gly
115 120 125

Ala Asp Gly Ser Leu Gln Pro Leu Gln Trp Ala Gly Gly Arg Pro Ala
130 135 140

Ala Pro Asp Glu Leu Arg Leu His Pro His Arg Pro Leu Leu Arg Gln
145 150 155 160

Ala Gly Val His Pro Gln Arg Arg Gly Gln Leu Gln Pro His Tyr Pro
165 170 175

Asp Phe Gly Arg Leu His His Arg Ser His Gly Gly Asp Leu Cys Pro
180 185 190

Ala His Arg Cys Gly Glu Gly Pro Ile Ser Gly Gln His Thr Pro Arg
195 200 205

Ala Ile Gln Glu Arg Gln Lys Ile Gln Arg Asp Tyr Gly Arg Leu Gln
210 215 220

Asn His Arg Gln Gly Gly Arg Ser Gln Gly Pro Tyr Glu Arg Asn Phe
225 230 235 240

Ala Gln His His Glu Glu Cys Tyr Arg Gln Leu Gly Gly Gly Asp Leu
245 250 255

Arg His Pro Gln Gly Glu Ala Ala Gly Leu Pro Pro Ala His Gln Leu
260 265 270

Pro Leu Pro Leu Cys Leu Cys Leu Trp Ser Arg Leu Leu Cys His Ser
275 280 285

Gly Gly Leu Pro Gly Gly Arg Gly Glu Asp Pro Val Tyr Glu Leu Thr
290 295 300

Ser Arg Pro Val Leu Gln Pro Pro Arg Leu Tyr Asp Lys Asp Gly Gly
305 310 315 320

Pro Gly Gly Pro His Ser Leu Leu Gln Gly Ile Tyr Thr Leu Leu Phe
325 330 335

Ala Phe Gly Ile Leu Glu Arg Gly Asp Val Arg Asn Leu Ala Ala Glu
340 345 350

Thr Gly Pro Asp Glu Ser Pro Asp Val Thr Gly Ile Thr Val Leu Asn
355 360 365

Lys Thr Arg Arg Pro Leu Val Ala Lys Val Ser Glu Thr Ser Glu Trp
370 375 380

Lys Lys Thr Val His Pro Arg Thr His Gly His Arg Pro Thr His
385 390 395

(2) INFORMATION FOR SEQ ID NO:30:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 339 amino acids
(B) TYPE: amino acid
(C) STRANDEDNESS: unknown
(D) TOPOLOGY: unknown

(ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:30:

Arg Gly His Pro Ile Pro Ala Ala Thr Ser Trp Asp Gly Ala Leu Gly
1 5 10 15

Ser Pro Cys Ala Ala Pro Ala Val Ala Gly Leu Thr Ala Pro Pro Leu
20 25 30

Ser Pro Gly Leu Trp Ser Ser Leu Ser Pro Trp Thr Ser Ser Arg Pro
35 40 45

Arg Asp Trp Ala Glu Pro Ser Arg Thr Met Val Gly Leu Lys Pro Ser
50 55 60

Asp Val Pro Pro Thr Met Ala Val Lys Phe Leu Gly Ala Gly Thr Ala
65 70 75 80

Ala Cys Phe Ala Glu Leu Val Thr Phe Pro Leu Asp Thr Ala Lys Val
85 90 95

Arg Leu Gln Ile Gln Gly Glu Asn Gln Ala Val Gln Thr Ala Arg Leu
100 105 110

Val Gln Tyr Arg Gly Val Leu Gly Thr Ile Leu Thr Met Val Arg Thr
115 120 125

Glu Gly Pro Cys Ser Pro Tyr Asn Gly Leu Val Ala Gly Leu Gln Arg
130 135 140

Gln Met Ser Phe Ala Ser Ile Arg Ile Gly Leu Tyr Asp Ser Val Lys
145 150 155 160

Gln Val Tyr Thr Pro Lys Gly Ala Asp Asn Ser Ser Leu Thr Thr Arg
165 170 175

Ile Leu Ala Gly Cys Thr Thr Gly Ala Met Ala Val Thr Cys Ala Gln
180 185 190

Pro Thr Asp Val Val Lys Val Arg Phe Gln Ala Ser Ile His Leu Gly
195 200 205

Pro Ser Arg Ser Asp Arg Lys Tyr Ser Gly Thr Met Asp Ala Tyr Arg
210 215 220

Thr Ile Ala Arg Phe Glu Gly Val Arg Gly Leu Trp Lys Gly Thr Leu
225 230 235 240

Pro Asn Ile Met Arg Asn Ala Ile Val Asn Cys Ala Glu Val Val Thr
245 250 255

Tyr Asp Ile Leu Lys Glu Lys Leu Asp Tyr His Leu Leu Thr Asp Asn
260 265 270

Phe Pro Cys His Phe Val Ser Ala Phe Gly Ala Gly Phe Cys Ala Thr
275 280 285

Val Val Ala Ser Pro Val Asp Val Val Lys Thr Arg Tyr Met Asn Ser
290 295 300

Pro Pro Gly Gln Tyr Phe Ser Pro Leu Asp Cys Met Ile Lys Met Val
305 310 315 320

Ala Gln Glu Gly Pro Thr Ala Phe Tyr Lys Gly Ala Ser Ser Cys Leu
325 330 335

Gln His Ser

(2) INFORMATION FOR SEQ ID NO:31:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 331 amino acids
(B) TYPE: amino acid
(C) STRANDEDNESS: unknown
(D) TOPOLOGY: unknown

(ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:31:

Gly Ala Ile Gln Ser Leu Leu Pro Pro Pro Gly Met Glu Pro Gly Ala
1 5 10 15

Pro Val Leu Pro Leu Pro Trp Gln Asp Ser Gln Pro His Arg Cys Ile
20 25 30

Glu Ala Gln Gly Cys Gly Ala Ala Ser Leu Leu Gly Pro Pro Leu Gly
35 40 45

Pro Lys Gly Thr Gly Gln Ser Leu Pro Gly Leu Trp Leu Asp Ser Leu
50 55 60

Gln Thr Cys Leu Pro Pro Trp Leu Ser Ser Trp Gly Gln Ala Gln Gln
65 70 75 80

Pro Val Leu Leu Asn Ser Leu Pro Phe His Trp Thr Gln Pro Arg Ser
85 90 95

Ala Cys Arg Ser Arg Gly Arg Ile Arg Arg Ser Arg Arg Pro Gly Ser
100 105 110

Cys Ser Thr Val Ala Cys Trp Ala Pro Ser Pro Trp Cys Gly Leu Arg
115 120 125

Val Pro Ala Ala Pro Thr Met Gly Trp Trp Pro Ala Cys Ser Ala Arg
130 135 140

Ala Ser Pro Pro Ser Ala Ser Ala Ser Met Thr Pro Ser Ser Arg Cys
145 150 155 160

```
Thr Pro Pro Lys Ala Arg Thr Thr Pro Ala Ser Leu Pro Gly Phe Trp
                165                 170                 175

Pro Ala Ala Pro Gln Glu Pro Trp Arg Pro Val Pro Ser Pro Gln Met
            180                 185                 190

Trp Arg Ser Asp Phe Arg Pro Ala Tyr Thr Ser Gly His Pro Gly Ala
        195                 200                 205

Thr Glu Asn Thr Ala Gly Leu Trp Thr Pro Thr Glu Pro Ser Pro Gly
    210                 215                 220

Arg Lys Glu Ser Gly Ala Cys Gly Lys Glu Leu Cys Pro Thr Ser Gly
225                 230                 235                 240

Met Leu Ser Ser Thr Val Leu Arg Trp Pro Thr Thr Ser Ser Arg Arg
                245                 250                 255

Ser Cys Trp Thr Thr Thr Cys Ser Leu Thr Thr Ser Pro Ala Thr Leu
            260                 265                 270

Ser Leu Pro Leu Glu Pro Ala Ser Val Pro Gln Trp Trp Pro Pro Arg
        275                 280                 285

Trp Ile Trp Arg Pro Gly Ile Thr His Leu Gln Ala Ser Thr Ser Ala
    290                 295                 300

Pro Ser Thr Val Arg Trp Trp Pro Arg Arg Ala Pro Gln Pro Ser Thr
305                 310                 315                 320

Arg Gly Glu Pro Pro Pro Ala Ser Ser Thr Pro
                325                 330
```

(2) INFORMATION FOR SEQ ID NO:32:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 340 amino acids
(B) TYPE: amino acid
(C) STRANDEDNESS: unknown
(D) TOPOLOGY: unknown

(ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:32:

```
Glu Gly Pro Ser Asn Pro Cys Cys His Leu Leu Gly Trp Ser Pro Arg
1               5                   10                  15

Glu Pro Leu Cys Cys Pro Cys Arg Gly Arg Thr His Ser Pro Thr Ala
            20                  25                  30

Ala Leu Lys Pro Arg Ala Val Glu Gln Pro Leu Ser Leu Asp Leu Leu
        35                  40                  45

Ser Ala Leu Lys Gly Leu Gly Arg Ala Phe Gln Asp Tyr Gly Trp Thr
    50                  55                  60
```

Glu Ala Phe Arg Arg Ala Ser His His Gly Cys Glu Val Pro Gly Gly
65 70 75 80

Arg His Ser Ser Leu Phe Cys Thr Arg Tyr Leu Ser Thr Gly His Ser
85 90 95

Gln Gly Pro Pro Ala Asp Pro Gly Gly Glu Pro Gly Gly Pro Asp Gly
100 105 110

Pro Ala Arg Ala Val Pro Trp Arg Ala Gly His His Pro Asp His Gly
115 120 125

Ala Asp Gly Ser Leu Gln Pro Leu Gln Trp Ala Gly Gly Arg Pro Ala
130 135 140

Ala Pro Asp Glu Leu Arg Leu His Pro His Arg Pro Leu Leu Arg Gln
145 150 155 160

Ala Gly Val His Pro Gln Arg Arg Gly Gln Leu Gln Pro His Tyr Pro
165 170 175

Asp Phe Gly Arg Leu His His Arg Ser His Gly Gly Asp Leu Cys Pro
180 185 190

Ala His Arg Cys Gly Glu Gly Pro Ile Ser Gly Gln His Thr Pro Arg
195 200 205

Ala Ile Gln Glu Arg Gln Lys Ile Gln Arg Asp Tyr Gly Arg Leu Gln
210 215 220

Asn His Arg Gln Gly Gly Arg Ser Gln Gly Pro Val Glu Arg Asn Phe
225 230 235 240

Ala Gln His His Glu Glu Cys Tyr Arg Gln Leu Cys Gly Gly Asp Leu
245 250 255

Arg His Pro Gln Gly Glu Ala Ala Gly Leu Pro Pro Ala His Cys Leu
260 265 270

Pro Leu Pro Leu Cys Leu Cys Leu Trp Ser Arg Leu Leu Cys His Ser
275 280 285

Gly Gly Leu Pro Gly Gly Arg Gly Glu Asp Pro Val Tyr Glu Leu Thr
290 295 300

Ser Arg Pro Val Leu Gln Pro Pro Arg Leu Tyr Asp Lys Asp Gly Gly
305 310 315 320

Pro Gly Gly Pro His Ser Leu Leu Cys Gly Val Ser Leu Leu Leu Pro
325 330 335

Pro Ala Leu Pro
340

(2) INFORMATION FOR SEQ ID NO:33:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 397 amino acids
(B) TYPE: amino acid
(C) STRANDEDNESS: unknown
(D) TOPOLOGY: unknown

(ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:33:

Glu Thr Thr Val Asn Gly Glu Ala Arg Pro Ser Asp Pro Ala Ala Thr
1 5 10 15

Trp Ser Cys Ala Ile Gly Trp Pro Cys Thr Thr Gln Pro Trp Leu Asp
20 25 30

Ala Gln Leu Pro Pro Thr Glu Ala Lys Asp Cys Gln Ala Ser Ser Leu
35 40 45

Leu Gly Pro Pro Ala Ala Lys Glu Pro Gly Pro Phe Pro Gly Thr Met
50 55 60

Val Gly Leu Gln Pro Ser Glu Val Pro Pro Thr Thr Val Val Lys Phe
65 70 75 80

Leu Gly Ala Gly Thr Ala Ala Cys Phe Ala Asp Leu Leu Thr Phe Pro
85 90 95

Leu Asp Thr Ala Lys Val Arg Leu Gln Ile Gln Gly Glu Asn Pro Gly
100 105 110

Ala Cys Ser Val Gln Tyr Arg Gly Val Leu Gly Thr Ile Leu Thr Met
115 120 125

Val Arg Thr Glu Gly Pro Arg Ser Pro Tyr Ser Gly Leu Val Ala Gly
130 135 140

Leu His Arg Gln Met Ser Phe Ala Ser Ile Arg Ile Gly Leu Tyr Asp
145 150 155 160

Ser Val Lys Gln Phe Tyr Thr Pro Lys Gly Ala Asp His Ser Ser Val
165 170 175

Ala Ile Arg Ile Leu Ala Gly Cys Thr Thr Gly Ala Met Ala Val Thr
180 185 190

Cys Ala Gln Pro Thr Asp Val Val Lys Val Arg Phe Gln Ala Met Ile
195 200 205

Arg Leu Gly Thr Gly Gly Glu Arg Lys Tyr Arg Gly Thr Met Asp Ala
210 215 220

Tyr Arg Thr Ile Ala Arg Glu Glu Gly Val Arg Gly Leu Trp Lys Gly
225 230 235 240

Thr Trp Pro Asn Ile Thr Arg Asn Ala Ile Val Asn Cys Ala Glu Met
245 250 255

Val Thr Tyr Asp Ile Ile Lys Glu Lys Leu Leu Glu Ser His Leu Phe
260 265 270

Thr Asp Asn Phe Pro Cys His Phe Val Ser Ala Phe Gly Ala Gly Phe
275 280 285

Cys Ala Thr Val Val Ala Ser Pro Val Asp Val Val Lys Thr Arg Tyr
290 295 300

```
Met Asn Ala Pro Leu Gly Arg Tyr Arg Ser Pro Leu His Cys Met Leu
305                 310                 315                 320

Lys Met Val Ala Gln Glu Gly Pro Thr Ala Phe Tyr Lys Gly Phe Val
                325                 330                 335

Pro Ser Phe Leu Arg Leu Gly Ala Trp Asn Val Met Met Phe Val Thr
            340                 345                 350

Tyr Glu Gln Leu Lys Arg Ala Leu Met Lys Val Gln Val Leu Arg Glu
        355                 360                 365

Ser Pro Phe Thr Arg Gln Ala Gly Cys Leu Glu Gln Asn Lys Ala Ser
    370                 375                 380

Leu Pro Gly Thr Gln Ala His Thr Ser Glu Pro Cys Thr
385                 390                 395
```

(2) INFORMATION FOR SEQ ID NO:34:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 381 amino acids
(B) TYPE: amino acid
(C) STRANDEDNESS: unknown
(D) TOPOLOGY: unknown

(ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:34:

```
Arg Gln Gln Met Val Arg Pro Gly Arg Gln Ile Leu Leu Leu Pro Asn
1               5                   10                  15

Gly Val Glu Pro Gly Gly Pro Ala Leu Pro Asn Leu Gly Thr His Ser
            20                  25                  30

Phe Leu Pro Glu Leu Lys Gln Lys Ile Ala Arg Gln Ala Leu Ser Ser
        35                  40                  45

Asp Leu His Arg Gln Gln Arg Asn Gln Ala His Ser Pro Gly Pro Trp
    50                  55                  60

Leu Asp Phe Ser Pro Pro Lys Cys Leu Pro Gln Arg Leu Ser Ser Trp
65                  70                  75                  80

Gly Pro Ala Leu Arg Pro Val Leu Arg Thr Ser Ser Leu Phe Pro Trp
                85                  90                  95

Thr Pro Pro Arg Ser Val Cys Arg Ser Lys Gly Arg Thr Gln Gly Leu
            100                 105                 110

Arg Ala Cys Ser Thr Ala Val Cys Trp Val Pro Ser Leu Trp Cys Ala
        115                 120                 125

Asp Arg Val Pro Ala Ala Pro Thr Ala Asp Trp Ser Leu Ala Cys Thr
    130                 135                 140

Ala Arg Val Leu Pro Pro Phe Glu Leu Ala Ser Thr Thr Ile Ser Ser
145                 150                 155                 160
```

Ser Ser Thr Pro Pro Arg Glu Arg Thr Thr Pro Ala Ser Pro Ser Gly
165 170 175

Phe Trp Gln Ala Ala Arg Gln Glu Pro Trp Gln Pro Ala Pro Ser Pro
180 185 190

Arg Met Trp Arg Ser Asp Phe Lys Pro Tyr Ala Trp Glu Leu Glu Glu
195 200 205

Arg Gly Asn Thr Glu Gly Leu Trp Met Pro Thr Glu Pro Ser Pro Gly
210 215 220

Arg Lys Glu Ser Gly Ala Cys Gly Lys Gly Leu Gly Pro Thr Ser Gln
225 230 235 240

Glu Met Pro Leu Ser Thr Val Leu Arg Trp Pro Thr Thr Ser Ser Arg
245 250 255

Arg Ser Cys Trp Ser Leu Thr Cys Leu Leu Thr Thr Ser Pro Val Thr
260 265 270

Leu Ser Leu Pro Leu Glu Leu Ala Ser Val Pro Gln Trp Trp Pro Pro
275 280 285

Arg Trp Met Trp Arg Pro Asp Thr Thr Leu Pro Ala Gly Thr Ala Ala
290 295 300

Leu Cys Thr Val Cys Arg Trp Trp Leu Arg Arg Asp Pro Arg Pro Ser
305 310 315 320

Thr Lys Asp Leu Cys Pro Pro Phe Cys Val Trp Glu Leu Gly Thr Cys
325 330 335

Leu His Met Ser Asn Arg Gly Pro Lys Ser Arg Tyr Cys Gly Asn Leu
340 345 350

Arg Phe Glu Gln Gly Lys Gln Ala Ala Trp Asn Arg Thr Lys Arg Leu
355 360 365

Cys Leu Gly His Arg Pro Thr Arg Gln Asn Arg Ala Arg
370 375 380

(2) INFORMATION FOR SEQ ID NO:35:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 397 amino acids
(B) TYPE: amino acid
(C) STRANDEDNESS: unknown
(D) TOPOLOGY: unknown

(ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:35:

Arg Asp Asn Ser Glu Trp Gly Pro Ala Val Arg Ser Cys Cys Tyr Leu
1 5 10 15

Met Glu Trp Ser Leu Arg Val Ala Leu His Tyr Pro Thr Leu Ala Arg
20 25 30

Arg Thr Ala Ser Ser Leu Asn Ser Lys Arg Leu Pro Gly Lys Leu Ser
35 40 45

Pro Arg Thr Ser Ile Gly Ser Lys Gly Thr Arg Pro Ile Pro Arg Asp
50 55 60

His Gly Trp Thr Ser Ala Leu Arg Ser Ala Ser His Asn Gly Cys Glu
65 70 75 80

Val Pro Gly Gly Arg His Cys Gly Leu Phe Cys Gly Pro Pro His Phe
85 90 95

Ser Pro Gly His Arg Gln Gly Pro Ser Ala Asp Pro Arg Gly Glu Pro
100 105 110

Arg Gly Ser Glu Arg Ala Val Pro Arg Cys Ala Gly Tyr His Pro Asp
115 120 125

Tyr Gly Ala His Arg Gly Ser Pro Gln Pro Leu Gln Arg Thr Gly Arg
130 135 140

Trp Pro Ala Pro Pro Asp Glu Phe Cys Leu Met Ser Asn Trp Pro Leu
145 150 155 160

Arg Leu Cys Gln Ala Val Leu His Pro Gln Gly Ser Gly Pro Leu Gln
165 170 175

Arg Arg His Gln Asp Ser Gly Arg Leu His Asp Arg Ser His Gly Ser
180 185 190

Asp Leu Arg Pro Ala His Gly Cys Gly Glu Gly Pro Ile Ser Ser His
195 200 205

Asp Thr Pro Gly Asn Trp Arg Arg Glu Glu Ile Gln Arg Asp Tyr Gly
210 215 220

Cys Leu Gln Asn His Arg Gln Gly Gly Arg Ser Gln Gly Pro Val Glu
225 230 235 240

Arg Asp Leu Ala Gln His His Lys Lys Cys His Cys Gln Leu Cys Asp
245 250 255

Gly Asp Leu Arg His His Gln Gly Glu Val Ala Gly Val Ser Pro Val
260 265 270

Tyr Gln Leu Pro Leu Ser Leu Cys Leu Cys Leu Trp Ser Trp Leu Leu
275 280 285

Cys His Ser Gly Gly Leu Pro Gly Gly Cys Gly Lys Asp Pro Ile His
290 295 300

Glu Arg Ser Pro Arg Gln Val Pro Gln Pro Ser Ala Leu Tyr Ala Glu
305 310 315 320

Asp Gly Gly Ser Gly Gly Thr His Gly Leu Leu Gln Arg Ile Cys Ala
325 330 335

Leu Leu Ser Ala Ser Gly Ser Leu Glu Arg Asp Asp Val Cys Asn Ile
340 345 350

Ala Thr Glu Glu Gly Leu Asn Glu Ser Pro Gly Thr Ala Gly Ile Ser
355 360 365

```
Val Leu Asn Lys Ala Ser Arg Leu Pro Gly Thr Glu Gln Ser Val Ser
    370                 375                 380

Ala Trp Asp Thr Gly Pro His Val Arg Thr Val His Ala
385                 390                 395
```

CLAIMS

What is claimed:

1. Isolated or recombinant nucleic acid which encodes a mammalian uncoupling protein 3.

2. The nucleic acid of Claim 1 wherein the uncoupling protein 3 is human.

3. The nucleic acid of Claim 1 selected from the group consisting of: SEQ ID NO: 1, SEQ ID NO: 2 and SEQ ID NO: 7.

4. The nucleic acid of Claim 1 wherein said nucleic acid hybridizes under stringent conditions with DNA selected from the group consisting of: SEQ ID NO: 1, the complement of SEQ ID NO:1, SEQ ID NO: 2 the complement of SEQ ID NO: 2, SEQ ID NO: 7 and the complement of SEQ ID NO: 7.

5. The nucleic acid of Claim 1 wherein the nucleic acid encodes an amino acid sequence selected from the group consisting of: SEQ ID NO: 3, SEQ ID NO: 4, and SEQ ID NO. 8.

6. A recombinant nucleic acid construct comprising the nucleic acid of Claim 1.

7. The recombinant nucleic acid construct of Claim 6 wherein the nucleic acid is selected from the group consisting of: SEQ ID NO: 1, SEQ ID NO: 2 and SEQ ID NO: 7.

8. The recombinant nucleic acid construct of Claim 6 wherein the nucleic acid encodes the amino acid sequence selected from the group consisting of: SEQ ID NO: 3, SEQ ID NO 4, and SEQ ID NO: 8.

9. The recombinant nucleic acid construct of Claim 6 wherein the nucleic acid is operably linked to an expression control sequence.

10. A host cell comprising the nucleic acid of Claim 1.

11. The host cell of Claim 10 wherein the nucleic acid is operably linked to an expression control sequence, whereby mammalian uncoupling protein 3 is expressed when the host cell is maintained under conditions suitable for expression.

12. A method for producing a mammalian uncoupling protein 3 comprising:

a) introducing into a host cell a nucleic acid construct comprising a nucleic acid which encodes a mammalian uncoupling protein 3; and

b) maintaining the host cells produced in step a) under conditions whereby the nucleic acid is expressed and the mammalian uncopling protein 3 is produced.

13. An antibody or functional portion thereof which binds mammalian uncoupling protein 3.

14. A method of detecting mammalian uncoupling protein 3 in a sample comprising:

a) contacting a sample with an antibody which binds uncoupling protein 3, under conditions suitable

for specific binding of said antibody to the mammalian uncoupling protein 3; and

b) detecting an antibody-mammalian uncoupling protein 3 complex,

wherein if the antibody-mammalian uncoupling protein complex is detected, mammalian uncoupling protein 3 is present in the sample.

15. A method of identifying an agent which alters uncoupling protein 3 activity comprising the steps of:

a) introducing into a host cell a nucleic acid construct comprising a nucleic acid which encodes a mammalian uncoupling protein 3;

b) maintaining the host cells produced in step a) under conditions appropriate for expression of the nucleic acid;

c) contacting the cells of b) with the agent; and

d) detecting mitochondrial electrical potential of the cells of c) in the presence of the agent,

wherein a change in mitochondrial electrical potential in the presence of the agent indicates that the agent alters uncoupling protein 3 activity.

16. The method of Claim 15 wherein the mitochondrial electrical potential is detected using fluorescence.

17. A method of identifying an agent which is an activator of uncoupling protien 3 activity comprising the steps of:

a) introducing into a host cell a nucleic acid construct comprising a nucleic acid which encodes a mammalian uncoupling protein 3;

b) maintaining the host cells produced in step a) under conditions appropriate for expression of the nucleic acid;

c) contacting the cells of b) with the agent; and
d) detecting mitochondrial electrical potential of the cells of c) in the presence of the agent;
wherein a reduction in mitochondrial electrical potential in the presence of the agent indicates that the agent is an activator uncoupling protein 3 activity.

18. The method of Claim 17 wherein the mitochondrial electrical potential is detected using fluorescence.

19. A method of identifying an agent which is an inhibitor of uncoupling protein 3 activity comprising the steps of:
a) introducing into a host cell a nucleic acid construct comprising a nucleic acid which encodes a mammalian uncoupling protein 3;
b) maintaining the host cells produced in step a) under conditions appropriate for expression of the nucleic acid;
c) contacting the cells of b) with the agent; and
d) detecting mitochondrial electrical potential of the cells in the presence of the agent;
wherein an increase in mitochondrial electrical potential in the presence of the agent indicates that the agent is an inhibitor uncoupling protein 3 activity.

20. The method of Claim 19 wherein the mitochondrial electrical potential is detected using fluorescence.

21. A method of inhibiting protein catabolism in a mammal comprising administering to the mammal an effective amount of an inhibitor of uncoupling protein 3.

22. A method of enhancing protein catabolism in a mammal comprising adminstering to the mammal an effective amount of an enhancer of uncoupling protein 3.

23. A method of inhibiting muscle wasting in a mammal comprising adminstering to the mammal an effective amount of an inhibitor of uncoupling protein 3.

24. Use of an inhibitor of uncoupling protein 3 in a method of inhibiting protein catabolism in a mammal, wherein the method comprises administering to the mammal an effective amount of an inhibitor of uncoupling protein 3.

25. Use of an enhancer of uncoupling protein 3 in a method of enhancing protein catabolism in a mammal, wherein the method comprises administering to the mammal an effective amount of an enhancer of uncoupling protein 3.

26. Use of an inhibitor of uncoupling protein 3 in a method of inhibiting protein catabolism in a mammal, wherein the method comprises administering to the mammal an effective amount of an inhibitor of uncoupling protein 3.

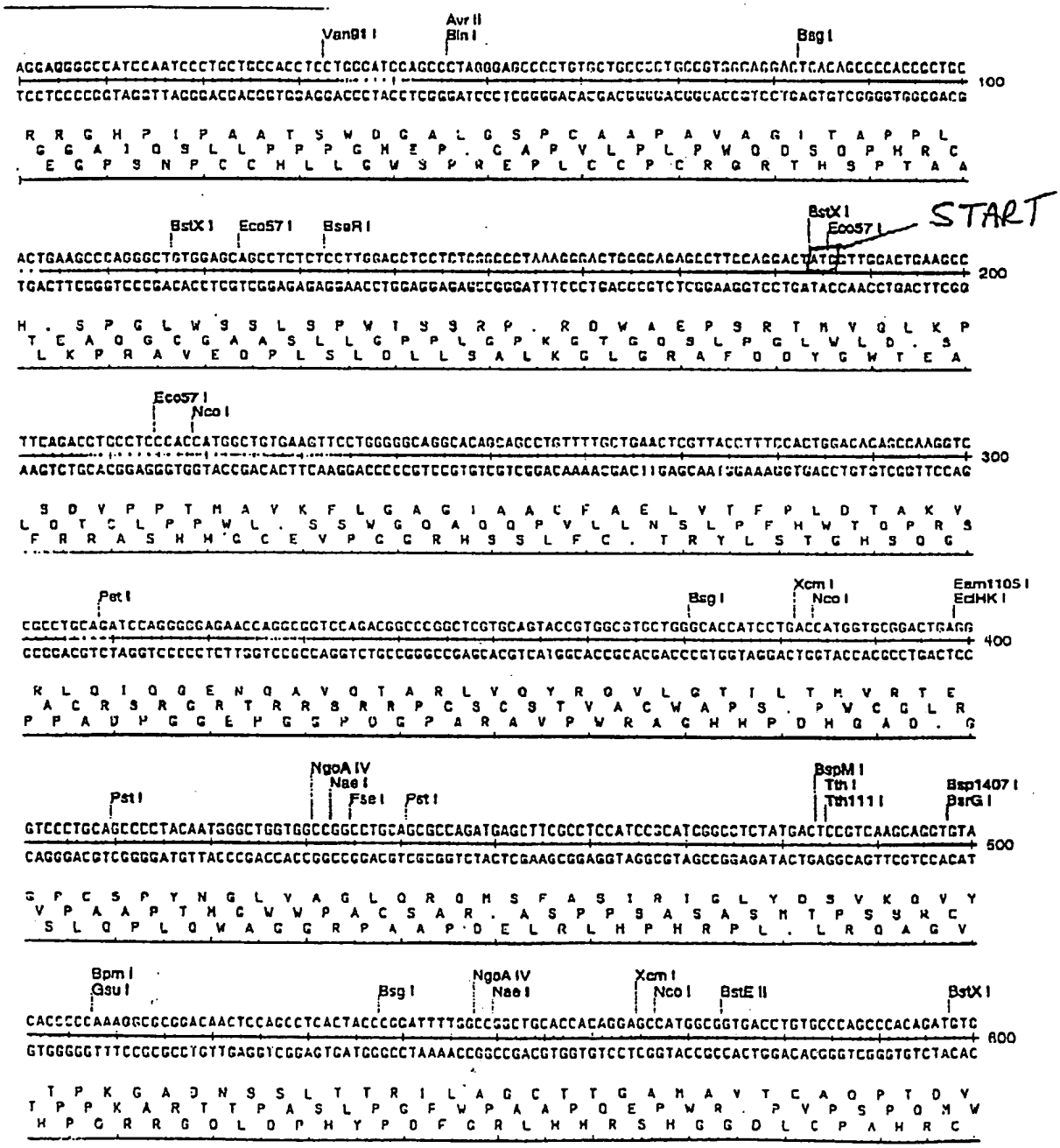
Van91 I
Avr II
Bln I
Bsg I
100
BstX I
Eco57 I
BsaR I
BstX I
Eco57 I
START
200
Eco57 I
Nco I
300
Pst I
Bsg I
Xcm I
Nco I
Eam1105 I
EclHK I
400
NgoA IV
Nae I
Fse I
Pst I
BspM I
Tth I
Tth111 I
Bsp1407 I
BsrG I
500
Bpm I
Gsu I
Bsg I
NgoA IV
Nae I
Xcm I
Nco I
BstE II
BstX I
600

FIGURE 1A

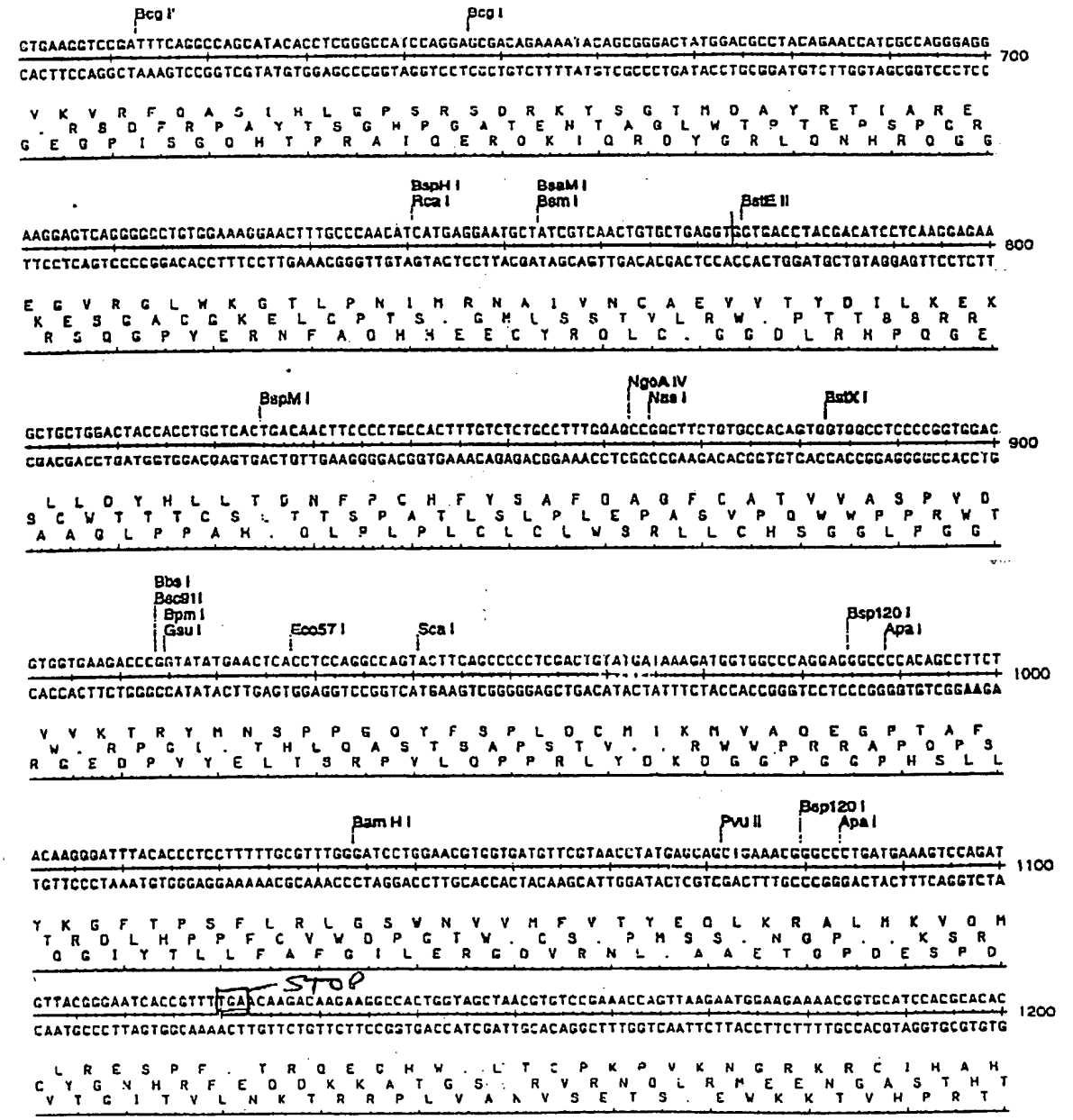
BcgI
BcgI
700
BspHI
RcaI
BsaMI
BsmI
BstEII
800
NgoAIV
NaeI
BstXI
900
BspMI
BbsI
Bsc91I
BpmI
GsuI
Eco57I
ScaI
Bsp120I
ApaI
1000
BamHI
PvuII
Bsp120I
ApaI
1100
STOP
1200

FIGURE 1B

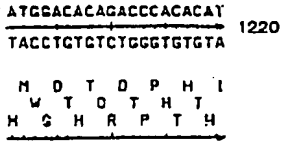
ATGGACACAGACCCACACAT
1220
TACCTGTGTCTGGGTGTGTA
M D T D P H I
W T Q T H T
H G H R P T H

FIGURE 1C

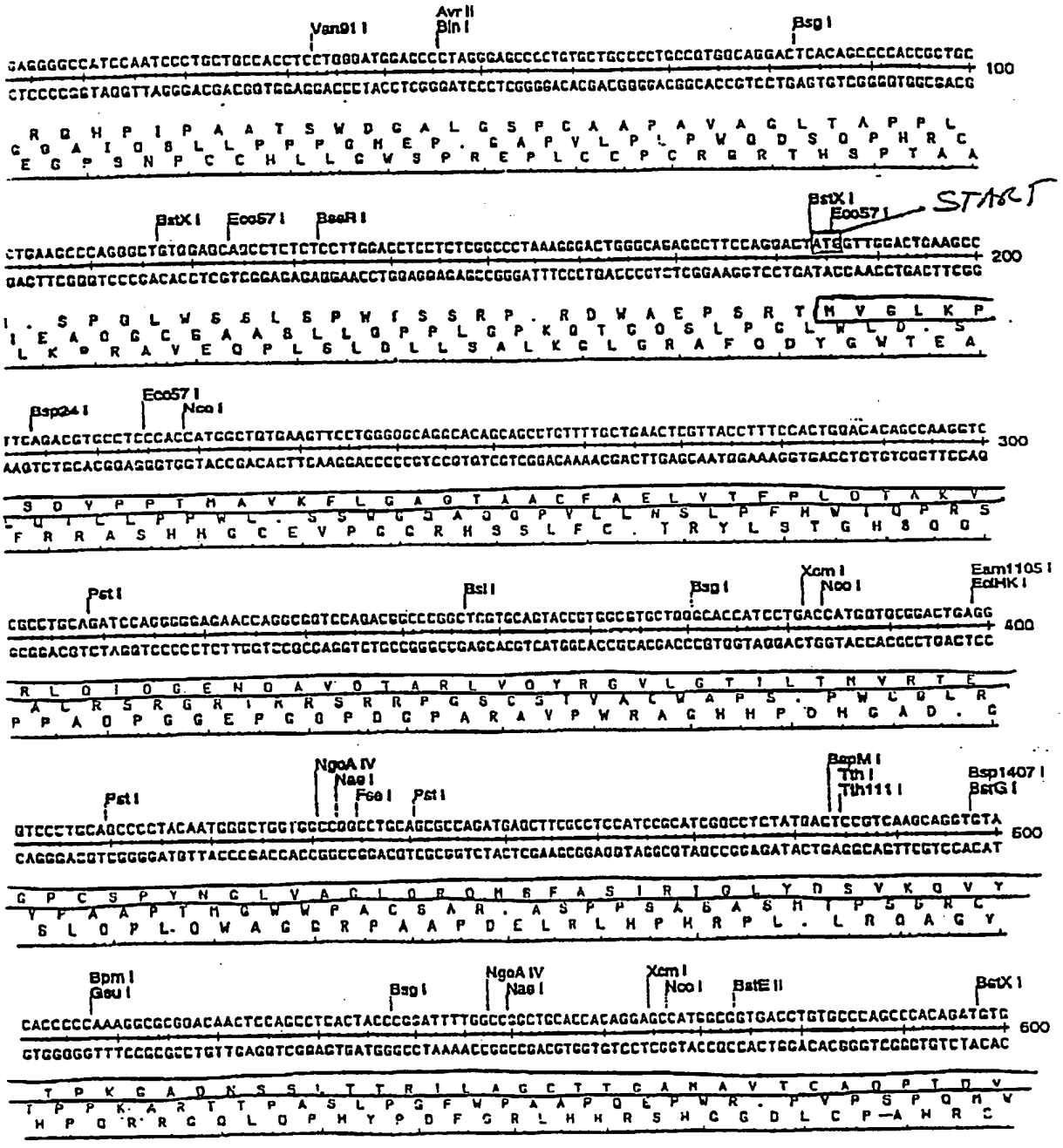
Van91 I
Avr II
Bln I
Bsg I
100
BstX I
Eco57 I
BseR I
START
200
Bsp24 I
Eco57 I
Nco I
300
Pst I
BsI I
Bsg I
Xcm I
Nco I
Esm1105 I
EclHK I
400
NgoA IV
Nae I
Fse I
Pst I
BspM I
Tth I
Tth111 I
Bsp1407 I
BstG I
500
Bpm I
Geu I
Bsg I
NgoA IV
Nae I
Xcm I
Nco I
BstE II
BstX I
600

Figure 2A

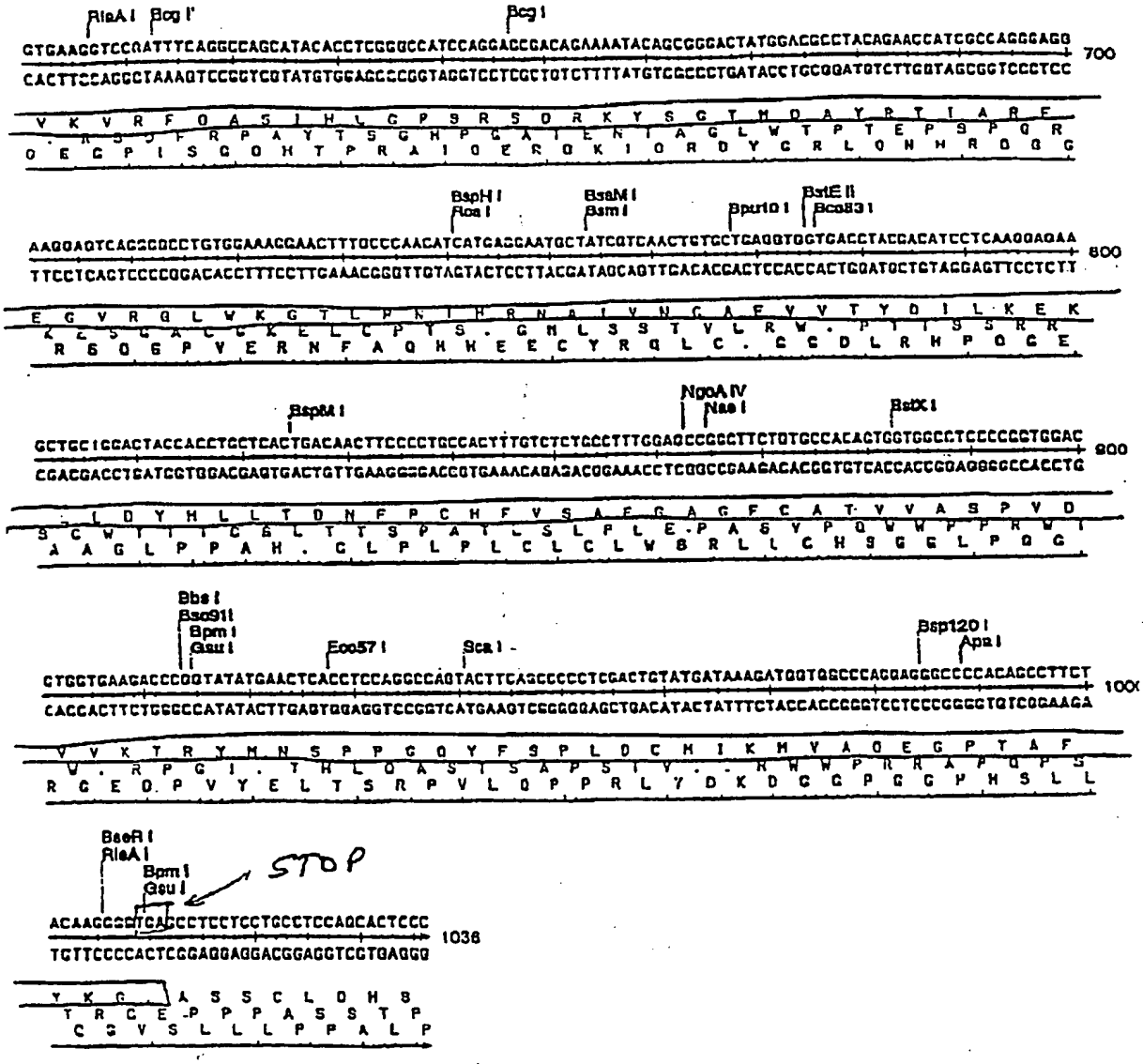
RleA I
Bcg I
Bcg I
700
BspH I
Roa I
BsaM I
Bsm I
Bpu10 I
BstE II
Bco83 I
800
NgoA IV
Nae I
BspM I
BstX I
900
Bbs I
Bso91I
Bpm I
Gsu I
Eco57 I
Sca I
Bsp120 I
Apa I
1000
BseR I
RleA I
Bpm I
Gsu I
STOP
1038

Figure 2B

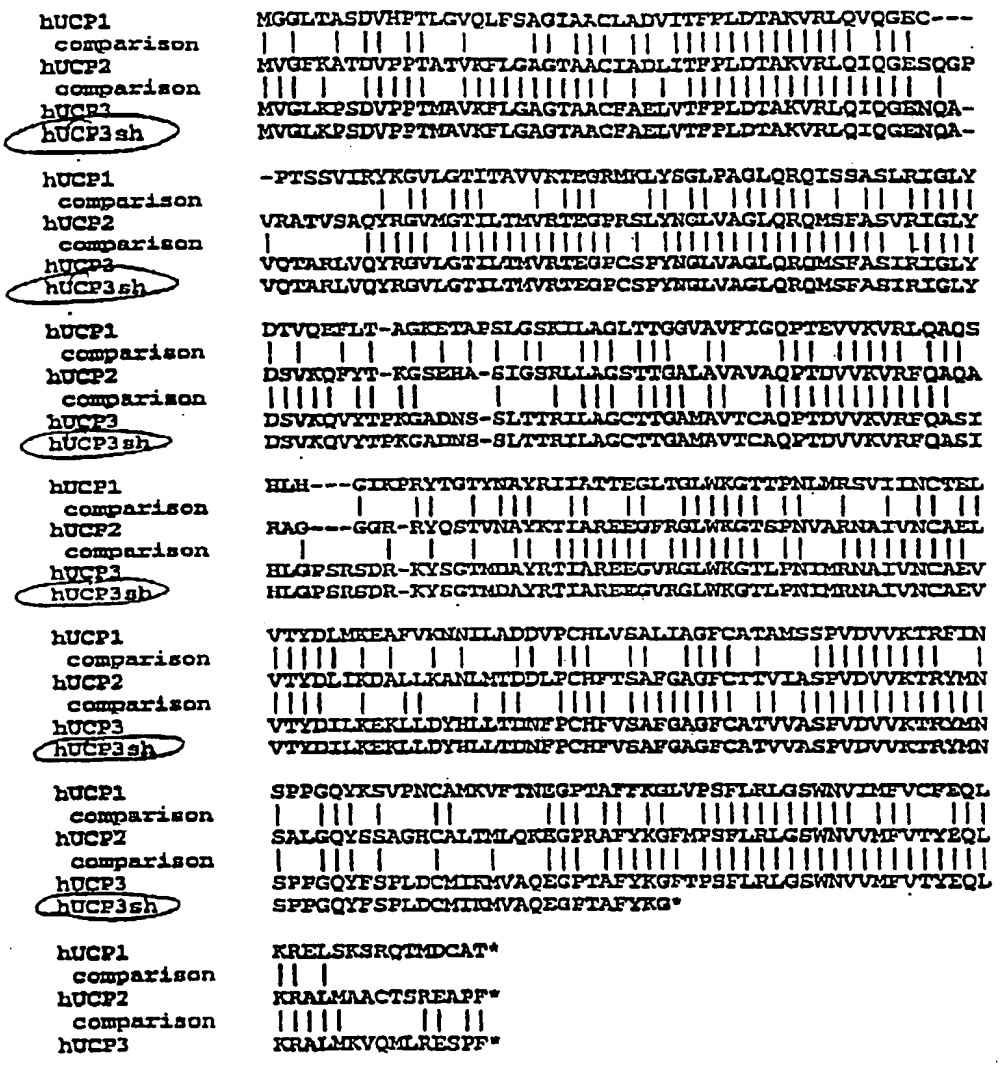
hUCP1
comparison
hUCP2
comparison
hUCP3
hUCP3sh
MGGLTASDVHPTLGVQLFSAGIAACLADVITFPLDTAKVRLQVQGEC---
MVGFKATDVPPTATVKFLGAGTAACIADLITFPLDTAKVRLQIQGESQGP
MVGLKPSDVPPTMAVKFLGAGTAACFAELVTFPLDTAKVRLQIQGENQA-
-PTSSVIRYKGVLGTITAVVKTEGRMKLYSGLPAGLQRQISSASLRIGLY
VRATVSAQYRGVMGTILTMVRTEGPRSLYNGLVAGLQRQMSFASVRIGLY
VQTARLVQYRGVLGTILTMVRTEGPCSFYNGLVAGLQRQMSFASIRIGLY
DTVQEFLT-AGKETAPSLGSKILAGLTTGGVAVFIGQPTEVVKVRLQAQS
DSVKQFYT-KGSEHA-SIGSRLLAGSTTGALAVAVAQPTDVVKVRFQAQA
DSVKQVYTPKGADNS-SLTTRILAGCTTGAMAVTCAQPTDVVKVRFQASI
HLH---GIKPRYTGTYNAYRIIATTEGLTGLWKGTTPNLMRSVIINCTEL
RAG---GGR-RYQSTVNAYKTIAREEGFRGLWKGTSPNVARNAIVNCAEL
HLGPSRSDR-KYSGTMDAYRTIAREEGVRGLWKGTLPNIMRNAIVNCAEV
VTYDLMKEAFVKNNILADDVPCHLVSALIAGFCATAMSSPVDVVKTRFIN
VTYDLIKDALLKANLMTDDLPCHFTSAFGAGFCTTVIASPVDVVKTRYMN
VTYDILKEKLLDYHLLTDNFPCHFVSAFGAGFCATVVASPVDVVKTRYMN
SPPGQYKSVPNCAMKVFTNEGPTAFYKGLVPSFLRLGSWNVIMFVCFEQL
SALGQYSSAGHCALTMLQKEGPRAFYKGFMPSFLRLGSWNVVMFVTYEQL
SPPGQYFSPLDCMIKMVAQEGPTAFYKGFTPSFLRLGSWNVVMFVTYEQL
SPPGQYPSPLDCMIKMVAQEGPTAFYKG*
KRELSKSRQTMDCAT*
KRALMAACTSREAPF*
KRALMKVQMLRESPF*

Figure 3

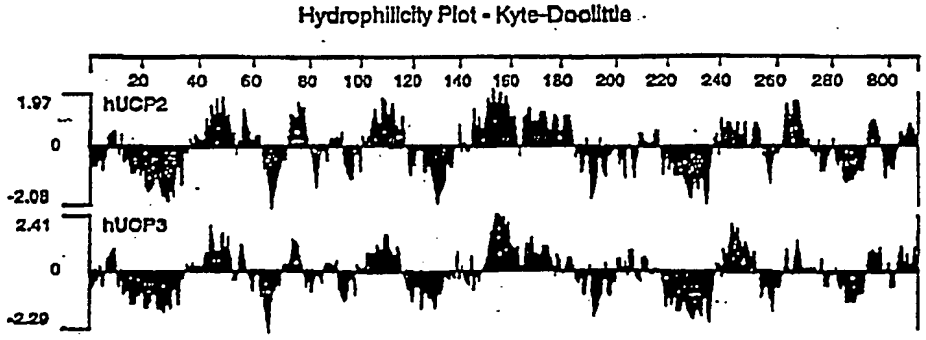
Hydrophilicity Plot - Kyte-Doolittle
20 40 60 80 100 120 140 160 180 200 220 240 260 280 300
1.97
hUCP2
0
-2.08
2.41
hUCP3
0
-2.29

Figure 4

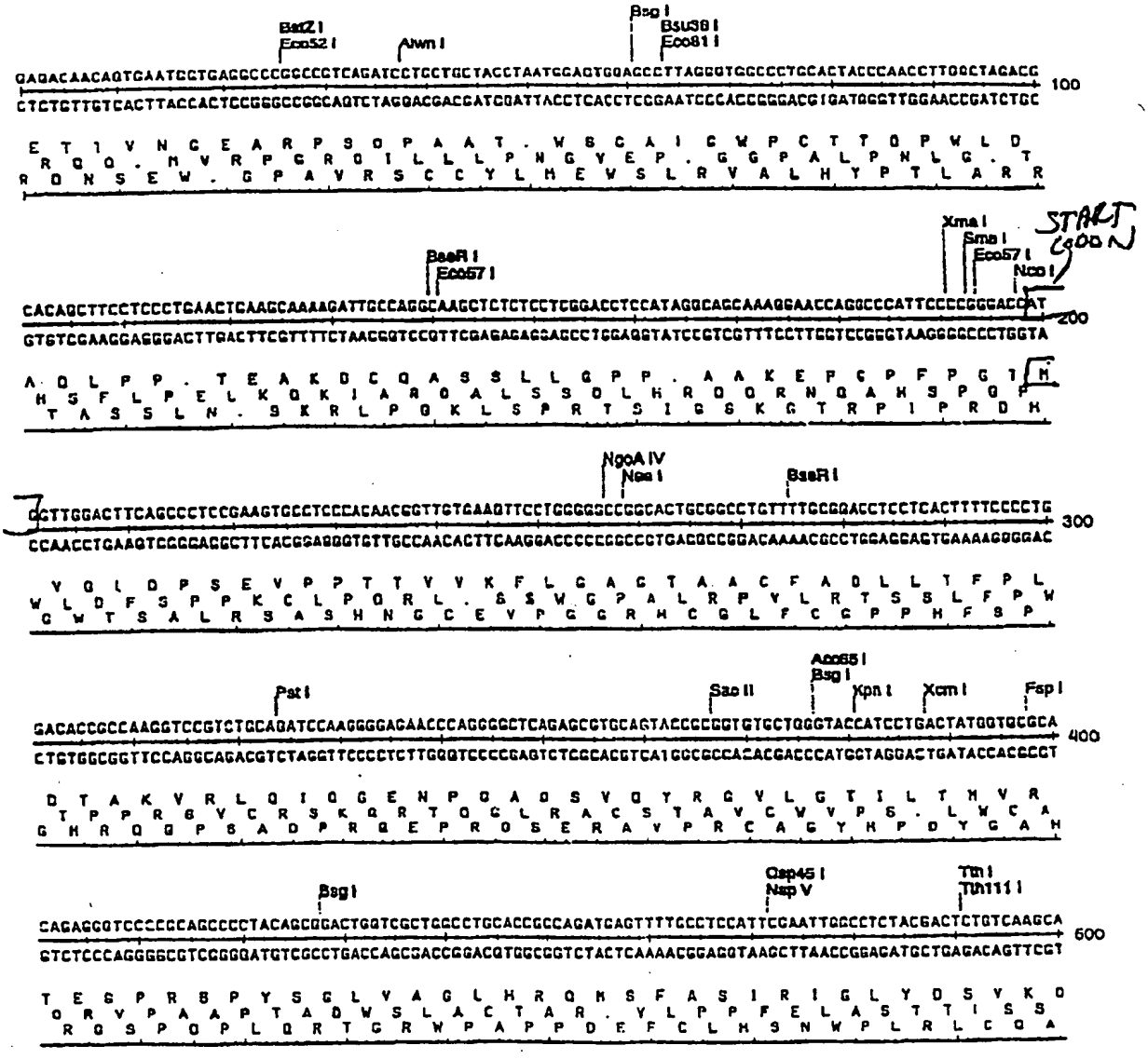
BstZ I
Eco52 I
Nwn I
Bsg I
BsiS8 I
Eco81 I
100
E T I V N C E A R P S Q P A A T . W S C A I C W P C T T Q P W L D
R Q Q . M V R P G R Q I L L L P N G Y E P . G G P A L P N L G . T
R Q N S E W . G P A V R S C C Y L M E W S L R V A L H Y P T L A R R
Xma I
Sma I
Eco57 I
Nco I
START
GGDN
BseR I
Eco57 I
200
A Q L P P . T E A K D C Q A S S L L Q P P . A A K E P G P F P G T M
H S F L P E L K Q K I A R Q A L S S Q L H R Q Q R N Q A H S P Q P
T A S S L N . S K R L P Q K L S P R T S I G S K G T R P I P R D H
NgoA IV
Nae I
BseR I
300
V Q L Q P S E V P P T T V V K F L G A G T A A C F A D L L T F P L
W L D F S P P K C L P Q R L . S S W G P A L R P Y L R T S S L F P W
G W T S A L R S A S H N G C E V P G G R H C G L F C G P P H F S P
Acc65 I
Bsg I
Pst I
Sac II
Xpn I
Xcm I
Fsp I
400
D T A K V R L Q I Q G E N P Q A Q S V Q Y R G V L G T I L T H V R
T P P R G V C R S K Q R T Q G L R A C S T A V C V V P S . L W C A
G H R Q Q P S A D P R Q E P R Q S E R A V P R C A G Y H P D Y G A H
Csp45 I
Nsp V
Tth
Tth111 I
Bsg I
600
T E S P R S P Y S G L V A G L H R Q H S F A S I R I G L Y D S V K Q
Q R V P A A P T A D W S L A C T A R . Y L P P F E L A S T I S S
R Q S P Q P L Q R T G R W P A P P D E F C L H S N W P L R L C Q A

Figure 5A

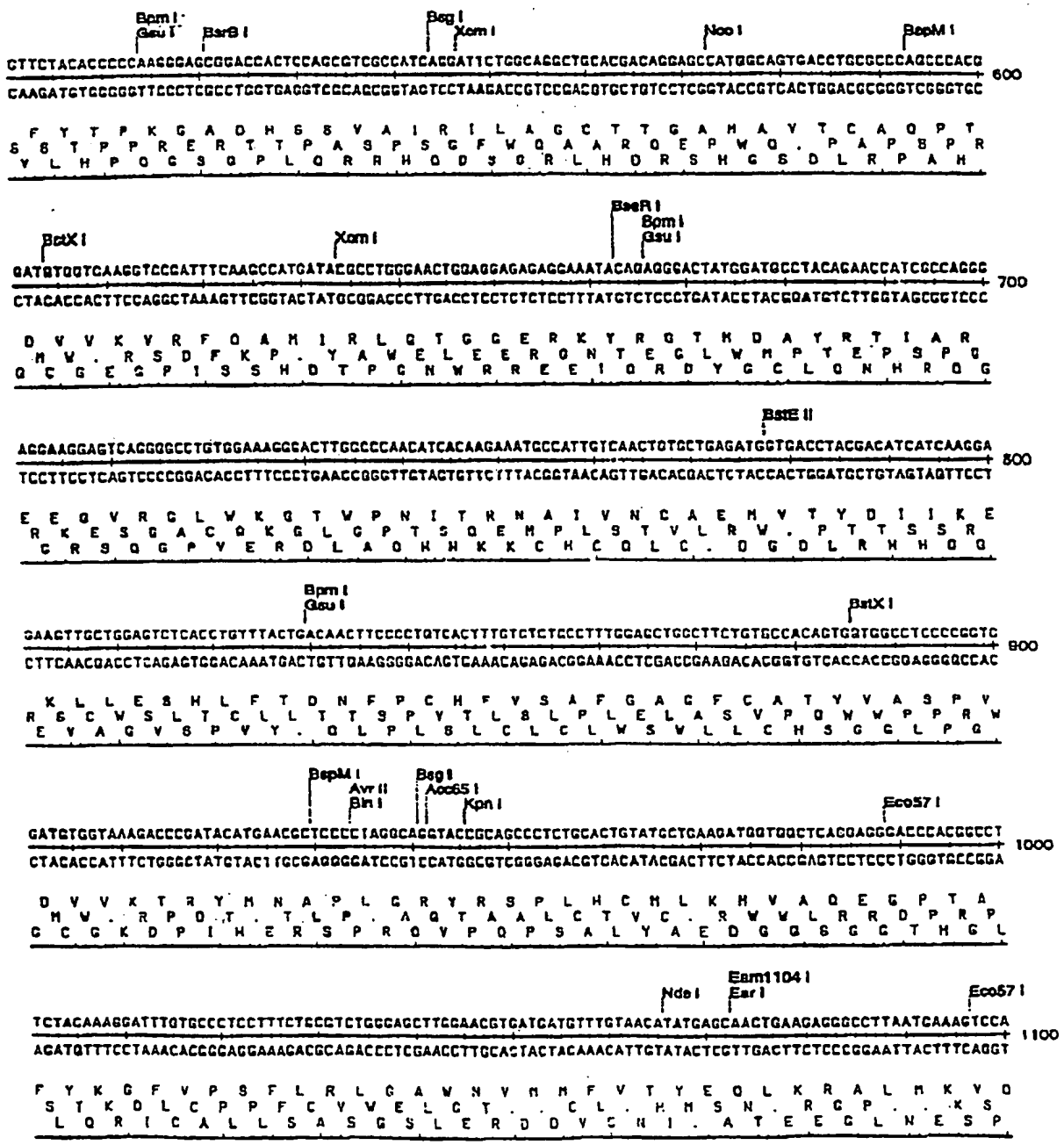
BpmI
GsuI
BsrBI
BsgI
XcmI
NcoI
BspMI
600
BstXI
XcmI
BseRI
BpmI
GsuI
700
BstEII
800
BpmI
GsuI
BstXI
900
BspMI
AvrII
BlnI
BsgI
Aoc65I
KpnI
Eco57I
1000
NdeI
EarI
Eam1104I
Eco57I
1100

Figure 5B

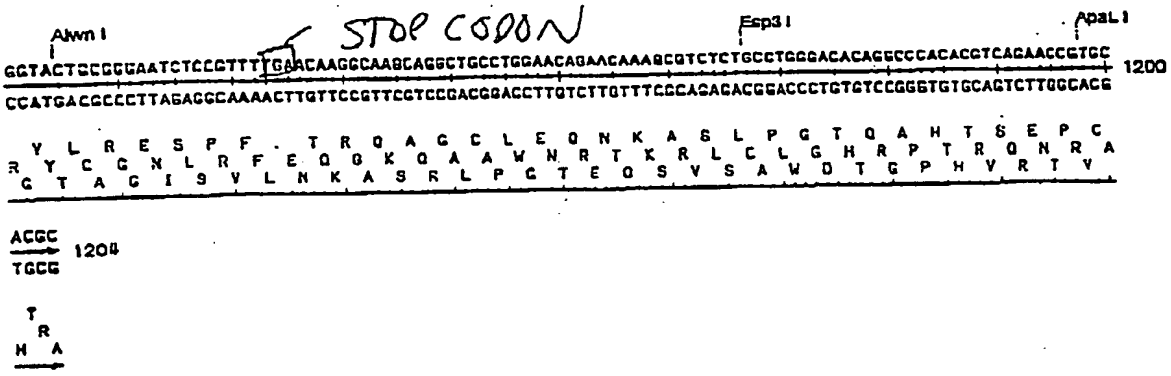

Figure 5C

---	Asp(D)	9	#	cua	Leu(L)	1	#	uca	Ser(S)	0	#	guu	Val(V)	2
ugc	Cys(C)	2	#	cuc	Leu(L)	3	#	ucc	Ser(S)	5	#	---	Val(V)	30
ugu	Cys(C)	5	#	cug	Leu(L)	18	#	ucg	Ser(S)	0	#	nnn	???(X)	0
---	Cys(C)	7	#	cuu	Leu(L)	1	#	ucu	Ser(S)	4	#	TOTAL		309
caa	Gln(Q)	3	#	uua	Leu(L)	1	#	---	Ser(S)	15	#			

```
         10        20        30        40
         |....|....|....|....|....|....|....|....|
MVGLQPSEVPPTTVVKFLGAGTAACFADLLTFPLDTAKVR 40
LQIQGENPGAQSVQYRGVLGTILTMVRTEGPRSPYSGLVA 80
GLHRQMSFASIRIGLYDSVKQFYTPKGADHSSVAIRILAG 120
CTTGAMAVTCAQPTDVVKVRFQAMIRLGTGGERKYRGTMD 160
AYRTIAREEGVRGLWKGTWPNITRNAIVNCAEMVTYDIIK 200
        210       220       230       240
        |....|....|....|....|....|....|....|
EKLLESHLFTDNFPCHFVSAFGAGFCATVVASPVDVVKTR 240
YMNAPLGRYRSPLHCMLKMVAQEGPTAFYKGFVPSFLRLG 280
AWNVMMFVTYEQLKRALMKVQVLRESPF. 309
```

Figure 6.

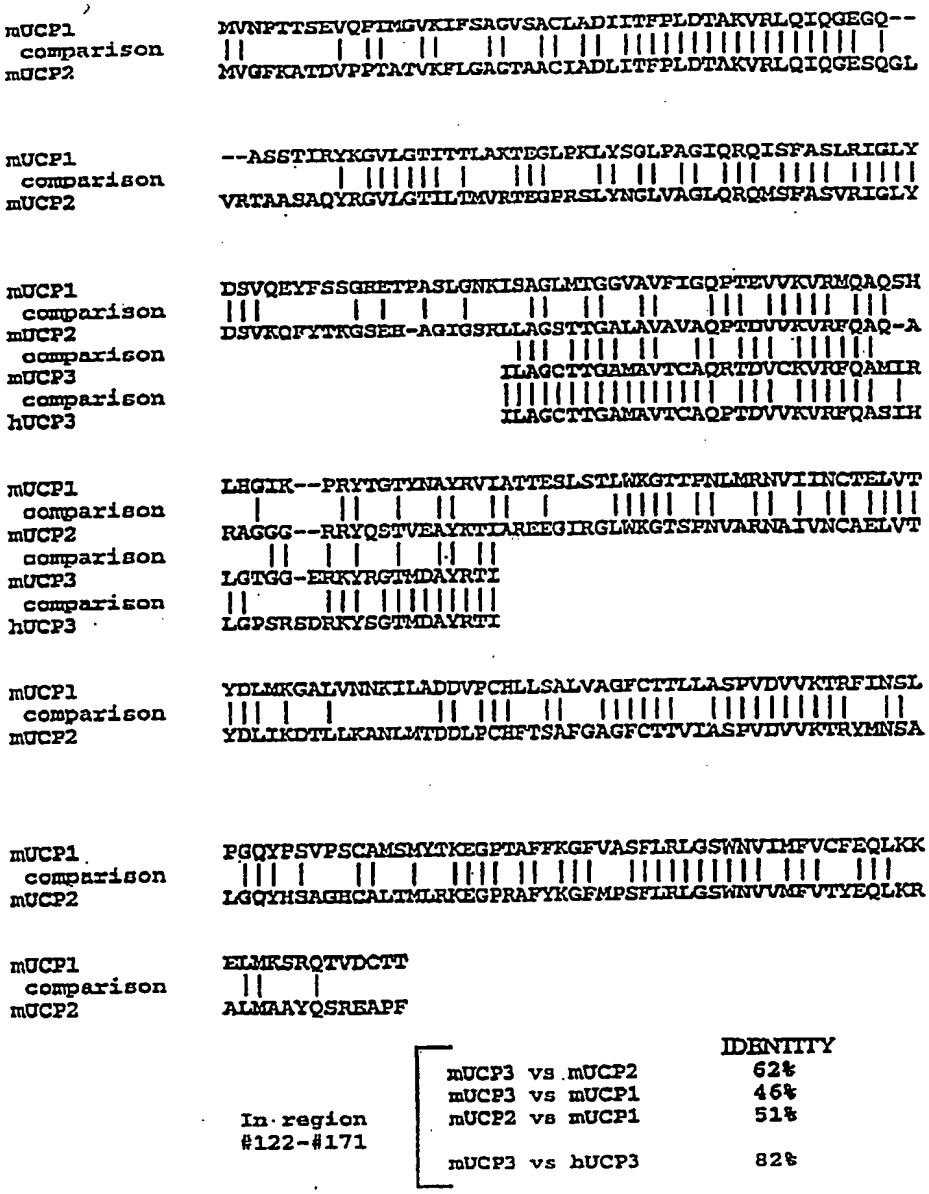
mUCP1
comparison
mUCP2
MVNPTTSEVQPIMGVKIFSAGVSACLADIITFPLDTAKVRLQIQGEGQ--
MVGFKATDVPPTATVKFLGAGTAACIADLITFPLDTAKVRLQIQGESQGL
--ASSTIRYKGVLGTITTLAKTEGLPKLYSGLPAGIQRQISFASLRIGLY
VRTAASAQYRGVLGTILTMVRTEGPRSLYNGLVAGLQRQMSFASVRIGLY
DSVQEYFSSGHETPASLGNKISAGLMTGGVAVFIGQPTEVVKVRMQAQSH
DSVKQFYTKGSEH-AGIGSKLLAGSTTGALAVAVAQPTDVVKVRFQAQ-A
ILAGCTTGAMAVTCAQRTDVCKVRFQAMIR
ILAGCTTGAMAVTCAQPTDVVKVRFQASIH
LHGIK--PRYTGTYNAYRVIATTESLSTLWKGTTPNLMRNVIINCTELVT
RAGGG--RRYQSTVEAYRTIAREEGIRGLWKGTSPNVARNAIVNCAELVT
LGTGG-ERKYRGTMDAYRTI
LGPSRSDRKYSGTMDAYRTI
YDLMKGALVNNKILADDVPCHLLSALVAGFCTTLLASPVDVVKTRFINSL
YDLIKDTLLKANLMTDDLPCHFTSAFGAGFCTTVIASPVDVVKTRYMNSA
PGQYPSVPSCAMSMYTKEGPTAFFKGFVASFLRLGSWNVIMFVCFEQLKK
LGQYHSAGHCALTMLRKEGPRAFYKGFMPSFLRLGSWNVVMFVTYEQLKR
ELMKSRQTVDCTT
ALMAAYQSREAPF
In region
#122-#171
IDENTITY
mUCP3 vs mUCP2 62%
mUCP3 vs mUCP1 46%
mUCP2 vs mUCP1 51%
mUCP3 vs hUCP3 82%

Figure 7

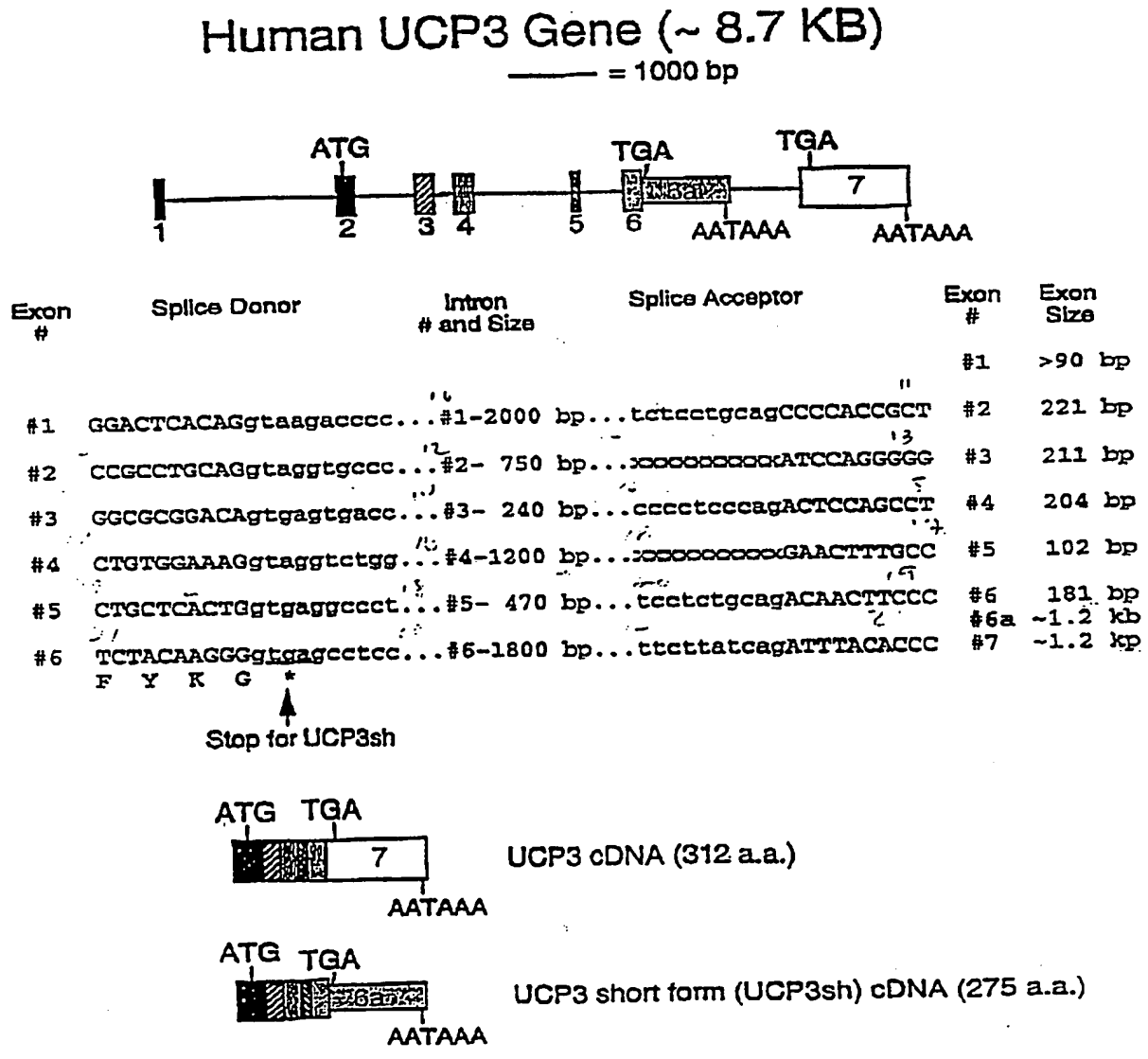
Human UCP3 Gene (~ 8.7 KB)
= 1000 bp
ATG
TGA
TGA
7
1
2
3
4
5
6
AATAAA
AATAAA
Exon #
Splice Donor
Intron # and Size
Splice Acceptor
Exon #
Exon Size
#1 >90 bp
#1 GGACTCACAGgtaagacccc...#1-2000 bp...tctcctgcagCCCCACCGCT #2 221 bp
#2 CCGCCTGCAGgtaggtgccc...#2- 750 bp...xxxxxxxxxxATCCAGGGGG #3 211 bp
#3 GGCGCGGACAgtgagtgacc...#3- 240 bp...ccctccccagACTCCAGCCT #4 204 bp
#4 CTGTGGAAAGgtaggtctgg...#4-1200 bp...xxxxxxxxxxGAACTTTGCC #5 102 bp
#5 CTGCTCACTGgtgaggccct...#5- 470 bp...tcctctgcagACAACTTCCC #6 181 bp
#6a ~1.2 kb
#6 TCTACAAGGGgtgagcctcc...#6-1800 bp...ttcttatcagATTTACACCC #7 ~1.2 kp
F Y K G *
Stop for UCP3sh
ATG TGA
7
AATAAA
UCP3 cDNA (312 a.a.)
ATG TGA
AATAAA
UCP3 short form (UCP3sh) cDNA (275 a.a.)

Figure 8

INTERNATIONAL SEARCH REPORT

rnational Application No
PCT/US 98/06959

A. CLASSIFICATION OF SUBJECT MATTER
IPC 6 C12N15/12 C07K14/705 C07K16/28 G01N33/50 A61K38/17
C12N1/21

According to International Patent Classification (IPC) or to both national classification and IPC

B. FIELDS SEARCHED

Minimum documentation searched (classification system followed by classification symbols)
IPC 6 C12N C07K G01N A61K

Documentation searched other than minimum documentation to the extent that such documents are included in the fields searched

Electronic data base consulted during the international search (name of data base and, where practical, search terms used)

C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category *	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
P,X	BOSS O ET AL: "Uncoupling protein-3: a new member of the mitochondrial carrier family with tissue-specific expression." FEBS LETT, MAY 12 1997, 408 (1) P39-42, XP002067895 NETHERLANDS see the whole document ---	1-8
P,X	VIDAL-PUIG A ET AL: "UCP3: an uncoupling protein homologue expressed preferentially and abundantly in skeletal muscle and brown adipose tissue." BIOCHEM BIOPHYS RES COMMUN, JUN 9 1997, 235 (1) P79-82, XP002075964 UNITED STATES see the whole document ---	1-8

-/--

[X] Further documents are listed in the continuation of box C.

[X] Patent family members are listed in annex.

* Special categories of cited documents :

"A" document defining the general state of the art which is not considered to be of particular relevance

"E" earlier document but published on or after the international filing date

"L" document which may throw doubts on priority claim(s) or which is cited to establish the publication date of another citation or other special reason (as specified)

"O" document referring to an oral disclosure, use, exhibition or other means

"P" document published prior to the international filing date but later than the priority date claimed

"T" later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention

"X" document of particular relevance; the claimed invention cannot be considered novel or cannot be considered to involve an inventive step when the document is taken alone

"Y" document of particular relevance; the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art.

"&" document member of the same patent family

Date of the actual completion of the international search	Date of mailing of the international search report
31 August 1998	14/09/1998

Name and mailing address of the ISA	Authorized officer
European Patent Office, P.B. 5818 Patentlaan 2 NL - 2280 HV Rijswijk Tel. (+31-70) 340-2040, Tx. 31 651 epo nl, Fax: (+31-70) 340-3016	Espen, J

Form PCT/ISA/210 (second sheet) (July 1992)

INTERNATIONAL SEARCH REPORT

rnational Application No
PCT/US 98/06959

C.(Continuation) DOCUMENTS CONSIDERED TO BE RELEVANT

Category	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
P,X	DA-WEI GONG ET AL: "Uncoupling protein-3 is a mediator of thermogenesis regulated by thyroid hormone, beta3-adrenergic agonists, and leptin" JOURNAL OF BIOLOGICAL CHEMISTRY, vol. 272, no. 39, 26 September 1997, pages 24129-24132, XP002075965 MD US see the whole document	1-12, 15-20
X	FLEURY C et AL: 'Human uncoupling protein-2 (UCP2) mRNA, nuclear gene encoding mitochondrial protein, complete CDS' EMHUM Database entry HSU76367 Accession number U76367; 06-MAR-1997 XP002075966	4
Y	see sequence	1-3,5
X	HILLIER L ET AL: 'Homo sapiens cDNA clone 628529 5' similar to TR:G412267 UNCOUPLING PROTEIN' EMEST Database entry Hsaa98452 Accession number AA192136; 21-01-1997 XP002075967	4
Y	see sequence	1-3,5
X	MARRA M ET AL: 'Mus musculus cDNA clone 570531 5' similar to SW:UCP_RABIT P14271 MITOCHONDRIAL BRWON FAT UNCOUPLING PROTEIN' EMEST Datbase entry Mmaa8362 Accession number AA108362; 06-NOV-1996 XP002075968	4
Y	see sequence	1-3,5
X	WO 96 05861 A (MILLENIUM PHARM INC) 29 February 1996	4
Y	see claims 1,3; figures 16,17	1-3,5

Form PCT/ISA/210 (continuation of second sheet) (July 1992)

INTERNATIONAL SEARCH REPORT

International application No.
PCT/US 98/06959

Box I Observations where certain claims were found unsearchable (Continuation of item 1 of first sheet)

This International Search Report has not been established in respect of certain claims under Article 17(2)(a) for the following reasons:

1. [X] Claims Nos.:
because they relate to subject matter not required to be searched by this Authority, namely:
Although claims 21-26 are directed to a method of treatment of the human/animal body, the search has been carried out and based on the alleged effects of the compounds/composition.

2. [ ] Claims Nos.:
because they relate to parts of the International Application that do not comply with the prescribed requirements to such an extent that no meaningful International Search can be carried out, specifically:

3. [ ] Claims Nos.:
because they are dependent claims and are not drafted in accordance with the second and third sentences of Rule 6.4(a).

Box II Observations where unity of invention is lacking (Continuation of item 2 of first sheet)

This International Searching Authority found multiple inventions in this international application, as follows:

1. [ ] As all required additional search fees were timely paid by the applicant, this International Search Report covers all searchable claims.

2. [ ] As all searchable claims could be searched without effort justifying an additional fee, this Authority did not invite payment of any additional fee.

3. [ ] As only some of the required additional search fees were timely paid by the applicant, this International Search Report covers only those claims for which fees were paid, specifically claims Nos.:

4. [ ] No required additional search fees were timely paid by the applicant. Consequently, this International Search Report is restricted to the invention first mentioned in the claims; it is covered by claims Nos.:

Remark on Protest

[ ] The additional search fees were accompanied by the applicant's protest.

[ ] No protest accompanied the payment of additional search fees.

INTERNATIONAL SEARCH REPORT

Information on patent family members

International Application No
PCT/US 98/06959

Patent document cited in search report	Publication date	Patent family member(s)	Publication date
WO 9605861 A	29-02-1996	US 5741666 A	21-04-1998
		AU 3497295 A	14-03-1996
		US 5702902 A	30-12-1997